

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 18, 2005, 20:14:54 ; Search time 192 Seconds

(without alignments)  
524.051 Million cell updates/sec

Title: US-09-942-052a-728

Sequence: 1 MAQPIRHRSRCATPPRGDP.....LKSIMKILSEVTDPDSKPEIN 229

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: _Geneseq_21.*
2: geneseqp1960s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003s:*
8: geneseqp2004s:*
9: geneseqp2005s:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	4	AAM40223 Human pol
2	1198	100.0	229	5	ABG34856 Human can
3	1198	100.0	231	4	AAM42009 Human pol
4	138.5	11.6	164	5	ABG34855 Human cdn
5	113	9.4	233	4	AAM93668 Human pol
6	113	9.4	233	4	AAG93290 Human pro
7	113	9.4	233	7	ADJ69163 Human hea
8	113	9.4	233	8	ADJ31517 Human pro
9	92	7.7	695	5	ABBS4167 Lactococc
10	92	7.7	695	8	ABBS29356 Bacteri
11	90	7.5	269	4	AAO08972 Human pol
12	89	7.4	572	8	ABM83354 Human dia
13	89	7.4	623	4	AAB93182 Human pro
14	89	7.4	623	5	ABBS97233 Novel hum
15	89	7.4	652	4	AAB93168 Human pro
16	89	7.4	664	4	AAB83843 Amino aci
17	89	7.4	664	4	AAB20219 Human che
18	88.5	7.4	388	8	ADQ08702 Ciona int
19	87.5	7.3	1245	7	ADC32818 Human nov
20	86.5	7.2	470	8	ABM83305 Human dia
21	86.5	7.2	669	6	ABU23878 Protein e
22	85	7.1	158	7	ADH88711 Enterococ
23	85.5	7.1	571	8	ADU07910 Amino aci
24	84.5	7.1	485	5	ABF51336 Human MDD

25	84	7.0	496	4	AAB95788 Human pro
26	84	7.0	496	7	ADC31196 Human nov
27	83.5	7.0	479	4	AAB396578 Propionib
28	83.5	7.0	479	6	ABM36197 Propionib
29	83	6.9	598	2	AAM59461 Microbial
30	83	6.9	598	2	AAM59463 Microbial
31	83	6.9	598	2	AAM59457 Microbial
32	83	6.9	598	2	AAM59458 Microbial
33	83	6.9	598	2	AAM59459 Microbial
34	83	6.9	598	2	AAM59462 Microbial
35	83	6.9	598	2	AAM59460 Microbial
36	83	6.9	598	2	AAM59464 Microbial
37	83	6.9	599	2	AAM59456 Microbial
38	83	6.9	606	1	APB90110 Polypepti
39	83	6.9	607	2	AAM59456 Microbial
40	83	6.9	607	2	AAM59465 Microbial
41	83	6.9	663	6	ABU23749 Protein e
42	82.5	6.9	278	8	ADX72944 Plant ful
43	82.5	6.9	325	3	AAV90242 Human cyc
44	82.5	6.9	493	8	ADU72405 Thale cre
45	82.5	6.9	493	9	AEAL6977 Arabidops

## ALIGNMENTS

RESULT 1  
AAM40223 standard; protein; 229 AA.  
ID AAM40223;

XX	AC	AAM40223;	Human polypeptide SEQ ID NO 3368.
XX	DT	22-OCT-2001 (first entry)	
XX	DE	Human polypeptide SEQ ID NO 3368.	
XX	XX	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;	
XX	KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
XX	KW	Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;	
XX	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
XX	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
XX	KW	leukemia.	
XX	OS	Homo sapiens.	
XX	PN	MO200153112-A1.	
XX	PD	26-JUL-2001.	
XX	PF	26-DEC-2000; 2000MO-US034263.	
XX	PR	23-DEC-1999; 99US-00471275.	
XX	PR	21-JAN-2000; 2000US-00488725.	
XX	PR	25-APR-2000; 2000US-00552317.	
XX	PR	20-JUN-2000; 2000US-00598042.	
XX	PR	19-JUL-2000; 2000US-00620312.	
XX	PR	03-AUG-2000; 2000US-00653450.	
XX	PR	14-SEP-2000; 2000US-00662191.	
XX	PR	19-OCT-2000; 2000US-00693036.	
XX	PR	29-NOV-2000; 2000US-00727344.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Qi;	
XX	PI	Wang J, Wang Z, Wehrman T, Xu C, xue AJ, Yang Y, Zhang J, Zhao J;	
XX	PI	Zhou P, Goodrich R, Drmanac RT,	
XX	DR	WPI, 2001-442253/47.	
XX	DR	N-PSDB; AAI59379.	
XX	PT	Novel nucleic acids and polypeptides, useful for treating disorders such	
XX	PT	as central nervous system injuries.	

PS Example 5; SEQ ID NO 3368; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with neurotrophic,  
CC immunosuppressant and cytostatic activity. The polypeptide or polynucleotide  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

XX Sequence 229 AA;

SQ

Query Match 100.0%; Score 1198; DB 4; Length 229;  
Best Local Similarity 100.0%; Pred. No. 9.7e-127;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPQLRRSRCAATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGAGAGGAE 60  
DB 1 MAAPQLRRSRCAATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGAGAGGAE 60  
QY 61 AAGPOLPSMLOPERCAVFCQAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120  
DB 61 AAGPOLPSMLOPERCAVFCQAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120  
QY 121 VGIEGSLKSTYNNLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKVCYLTKTAIV 180  
DB 121 VGIEGSLKSTYNNLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKVCYLTKTAIV 180  
QY 181 NASEMDIONVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229  
DB 181 NASEMDIONVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 2  
ABG34856  
ID ABG34856 standard; protein; 229 AA.  
XX  
AC ABG34856;  
XX  
DT 15-JUL-2002 (first entry)

DE Human cancer related protein encoded by cDNA 85PIB3.  
XX  
XX Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.  
XX  
OS Homo sapiens.  
XX  
PN WO200218578-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 28-AUG-2001; 2001WO-US026638.  
XX  
PR 28-AUG-2000; 2000US-0228432P.  
XX  
PA (AGEN-) AGENSYS INC.  
XX  
PI Raitano AB, Faris M, Hubert RS, Afar D, Ge W, Challita-Bid P;  
XX Jakovovits A;  
XX  
XX MPI; 2002-382963/41.  
XX  
XX DR N-PSDB; ABK70506.  
XX  
XX  
PT Composition for modulating the status of 85PIB3 protein or a molecule  
comprising a substance e.g. antibody specific to, nucleic acid encoding,

PT or ribozyme of 85PIB3.

XX  
XX  
PS Claim 34; Fig 2; 201pp; English.

XX The invention relates to a composition comprising a substance that  
CC modulate the status of 85PIB3, where the status of a cell expresses  
CC 85PIB3 gene product is modulated. Also included are a composition  
CC comprising a peptide region of 5 amino acids of the 85PIB3 protein, in  
CC any whole number increment up to 229 that includes an aa position  
CC selected from an aa position having a value greater than 0.5 in the  
CC Hydrophobicity profile, an aa position having a value less than 0.5 in  
CC the hydrophobicity profile, an aa position having a value greater than  
CC 0.5 in the percent accessible residue profile, an aa position having a  
CC value greater than 0.5 in the average flexibility profile, or an aa  
CC position having a value greater than 0.5 in the beta-turn profile; a  
CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous  
CC residues of the 85PIB3 protein; a recombinant protein comprising the  
CC antigen-binding region of a monoclonal antibody; a non-human transgenic  
CC animal that produces an antibody that binds to the 85PIB3 protein; a  
CC hybridoma that produces an antibody specific to the protein; a single chain  
CC monoclonal antibody (Mab) that comprises the variable domains of the  
CC heavy and monoclonal antibodies specific to the protein; a vector  
CC comprising a polynucleotide that encodes the Mab; inhibiting growth of  
CC cancer cells or treating a patient who bears cancer cells that expresses  
CC the protein, by administering the protein, antibody, polynucleotide  
CC encoding the protein, antisense polynucleotide to the polynucleotide,  
CC ribozyme that cleaves the polynucleotide and T cells that specifically  
CC recognize the protein; and generating a mammalian immune response  
CC to an immunogenic portion of the protein exposing cells of the mammal's immune system to  
CC an immunogenic portion of the protein or polynucleotide. The composition,  
CC which comprises an antibody specific to the protein, is useful for  
CC delivering a cytotoxic agent to a cell that expresses the protein by  
CC providing a cytotoxic agent conjugated to antibody and exposing the cell  
CC to the antibody-agent conjugate. The methods are useful for inhibiting  
CC growth of cancer cells or treating a patient who bears cancer cells that  
CC expresses the protein, for generating a mammalian immune response  
CC directed to the protein, for detecting the presence of the protein or  
CC polynucleotide in a biological sample in a patient who has or who is  
CC suspected of having cancer and for monitoring 85PIB3 in a biological  
CC sample from a patient who has or who is suspected of having cancer. The  
CC gene for 85PIB3 is located on human chromosome 15q14. The present  
CC sequence is the 85PIB3 protein

XX  
SQ Sequence 229 AA;

Query Match 100.0%; Score 1198; DB 5; Length 229;  
Best Local Similarity 100.0%; Pred. No. 9.7e-127;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPQLRRSRCAATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGAGAGGAE 60  
DB 1 MAAPQLRRSRCAATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGAGAGGAE 60  
QY 61 AAGPOLPSMLOPERCAVFCQAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120  
DB 61 AAGPOLPSMLOPERCAVFCQAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120  
QY 121 VGIEGSLKSTYNNLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKVCYLTKTAIV 180  
DB 121 VGIEGSLKSTYNNLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKVCYLTKTAIV 180  
QY 181 NASEMDIONVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229  
DB 181 NASEMDIONVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 3  
AAM42009  
ID AAM42009 standard; protein; 231 AA.  
XX  
XX  
XX AAM42009;  
XX  
XX  
XX  
DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6940.  
DE  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukemia.  
XX  
XX Homo sapiens.  
OS  
XX MO200153312-A1.  
PN  
XX 26-JUL-2001.  
PD  
XX 26-DEC-2000; 2000MO-US034263.  
PF  
XX 23-DEC-1999; 99US-00471275.  
PR  
XX 21-JAN-2000; 2000US-00488725.  
PR  
XX 25-APR-2000; 2000US-00552317.  
PR  
XX 20-JUN-2000; 2000US-00598042.  
PR  
XX 19-JUL-2000; 2000US-00620312.  
PR  
XX 03-AUG-2000; 2000US-00653450.  
PR  
XX 14-SEP-2000; 2000US-0062191.  
PR  
XX 19-OCT-2000; 2000US-00693036.  
PR  
XX 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;  
PI Zhou P, Goodrich R, Drmanac RT;  
PI  
XX WPI: 2001-442253/47.  
XX N-PSDB; AAI61165.  
DR  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
PT  
XX  
PS Example 2; SEQ ID NO 6940; 1007bpb; English.  
CC The invention relates to human nucleic acids (AA157798-AA161369) and the  
CC encoded polypeptides (AA08642-AA042213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localized neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilization of the activities such as: Immune system suppression,  
CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
XX Sequence 231 AA;  
SQ

Query Match 100.0%; Score 1198; DB 4; Length 231;  
Best Local Similarity 100.0%; Pred. No. 9.8e-127;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGPLHRRRCATPPGCGGCTGTERAIDQASFTSMEMPTQVVGSSPIGAGLGEPP 60  
DB 3 MAAGPLHRRRCATPPGCGGCTGTERAIDQASFTSMEMPTQVVGSSPIGAGLGEPP 62  
QY 61 AAGQLPSWLOPERCAVFOCAQCAHAVLADSVHLAMDLSRSIGAVVFSTRVNTNVLAPFL 120  
DB 63 AAGQLPSWLOPERCAVFOCAQCAHAVLADSVHLAMDLSRSIGAVVFSTRVNTNVLAPFL 122  
QY 121 VGEIGSLKSTYNNLLFCGSCGIPVGFHLVSTHAALALRGHFLCLSDKMCYLLKTKAIV 180

DB 123 VGEIGSLKSTYNNLLFCGSCGIPVGFHLVSTHAALALRGHFLCLSDKMCYLLKTKAIV 182  
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKLISVTPDQSKPEN 229  
DB 183 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKLISVTPDQSKPEN 231

RESULT 4  
ABG34855  
ID ABG34855 standard; protein; 164 AA.  
XX  
XX ABG34855;  
AC  
XX 15-JUL-2002 (first entry)  
DT  
XX Human cDNA 85PIB3 splice variant, open reading frame #3.  
DE  
XX Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.  
KW  
XX Homo sapiens.  
OS  
XX MO200218578-A2.  
PN  
XX 07-MAR-2002.  
PD  
XX 28-AUG-2001; 2001MO-US026838.  
PF  
XX 28-AUG-2000; 2000US-0228432P.  
PR  
XX  
XX (AGEN-) AGENSYS INC.  
PA  
XX Raitano AB, Faris M, Hubert RS, Afar D, Ge W, Challica-Bid P;  
PI Jakobovits A;  
PI  
XX WPI: 2002-382963/41.  
XX N-PSDB; ABR70504.  
DR  
XX  
XX Composition for modulating the status of 85PIB3 protein or a molecule  
PT comprising a substance e.g. antibody specific to, nucleic acid encoding,  
PT or ribozyme of 85PIB3.  
PT  
XX  
PS Example 38; Page 124; 201pb; English.  
CC The invention relates to a composition comprising a substance that  
CC modulate the status of 85PIB3, where the status of a cell expresses  
CC 85PIB3 gene product is modulated. Also included are a composition  
CC comprising a peptide region of 5 amino acids of the 85PIB3 protein, in  
CC any whole number increment up to 229 that includes an aa position  
CC selected from an aa position having a value greater than 0.5 in the  
CC Hydrophilicity profile, an aa position having a value less than 0.5 in  
CC the hydrophobicity profile, an aa position having a value greater than  
CC 0.5 in the percent accessible residue profile, an aa position having a  
CC value greater than 0.5 in the average flexibility profile, or an aa  
CC position having a value greater than 0.5 in the beta-turn profile; a  
CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous  
CC residues of the 85PIB3 protein; a recombinant protein comprising the  
CC antigen-binding region of a monoclonal antibody; a non-human transgenic  
CC animal that produces an antibody that binds to the 85PIB3 protein; a  
CC hybridoma that produces antibody specific to the protein; a single chain  
CC monoclonal antibody (MAb) that comprises the variable domains of the  
CC heavy and monoclonal antibodies specific to the protein; a vector  
CC comprising a polynucleotide that encodes the MAb; inhibiting growth of  
CC cancer cells or treating a patient who bears cancer cells that expresses  
CC the protein, by administering the protein, antibody, polynucleotide,  
CC encoding the protein, antisense polynucleotide and T cells that specifically  
CC recognize the protein; and generating a mammalian immune response  
CC directed to the protein exposing cells of the mammal's immune system to  
CC an immunogenic portion of the protein or polynucleotide. The composition,  
CC which comprises an antibody specific to the protein, is useful for  
CC delivering a cytotoxic agent to a cell that expresses the protein by  
CC providing a cytotoxic agent conjugated to antibody and exposing the cell

CC to the antibody-agent conjugate. The methods are useful for inhibiting  
 CC growth of cancer cells or treating a patient who bears cancer cells that  
 CC expresses the protein, for generating a mammalian immune response  
 CC directed to the protein, for detecting the presence of the protein or  
 CC polynucleotide in a biological sample in a patient who has or who is  
 CC suspected of having cancer and for monitoring 85PIB3 in a biological  
 CC sample from a patient who has or who is suspected of having cancer. The  
 CC gene for 85PIB3 is located on human chromosome 15q14. The present  
 CC sequence is the 85PIB3 protein  
 CC  
 XX

Sequence 164 AA;

Query Match 11.6%; Score 138.5; DB 5; Length 164;  
 Best Local Similarity 82.1%; Pred. No. 8.6e-07;  
 Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 162 FCLSSDKMVCYLLKTKAIVNASEMDIQNVPLSEKIAELK 200  
 Db 2 FFLSS-----YLLKTKAIVNASEMDIQNVPLSEKIAELK 35

RESULT 5  
 AAM93668  
 ID AAM93668 standard; protein; 233 AA.  
 XX  
 AC AAM93668;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide, SEQ ID NO: 3550.  
 XX  
 DE Human, full length cDNA; cDNA synthesis; oligo-capping.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI130094-A2.  
 XX  
 PD 05-SEP-2001.  
 XX  
 PF 07-JUL-2000; 2000EP-00114089.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 PR 11-JAN-2000; 2000JP-0018774.  
 PR 02-MAY-2000; 2000JP-00183765.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 DR WPI, 2001-524255/58.  
 DR N-PSDB; AAK94603.

830 Primers useful for synthesizing full length cDNA clones and their use  
 in genetic manipulation.

Claim 8; SEQ ID NO 3550; 1380bp + Sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been isolated  
 CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have  
 CC been determined. Primers for synthesizing the full length cDNA are useful  
 CC for clarifying the function of the protein encoded by the cDNA. The full  
 CC length clones were obtained by construction of full length enriched cDNA  
 CC libraries that were synthesized by the oligo-capping method. The primers  
 CC enable the production of the full length cDNA easily without any special  
 CC methods. The present sequence is a polypeptide encoded by a full length  
 CC human cDNA of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in CD-ROM  
 CC format directly from EPO  
 CC  
 XX Sequence 233 AA;  
 SQ

Query Match 9.4%; Score 113; DB 4; Length 233;  
 Best Local Similarity 27.1%; Pred. No. 0.0011;  
 Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

Qy 31 ASFTTSEMTQYVKGSPFPGAGLAEPAAGQPLSWIQPRCAVFCQAGCAVLADS 90  
 Db 49 ASWMSSESDASV---ADMERQQL--EEBAALAE-----ERPLVFLSGGCRPLGDS 95

Qy 91 VHLAMDLSR-SLGAIVPSRYTNVNLAPPLVGLGSLKSTYVILFCGSCGIPVGFHLY 149  
 Db 96 --LSWVASQEDPTNCILRCVSCNVSVDEKQLSKREKENCVLFTLCAGCSNLGVYR 153

Qy 150 STHAALALRGHFLCSDSKMVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207  
 Db 154 CTPGNLDLYKXKDLFCLSVLAIESVYLGSSSEKQIV-SEDKELFNL---ESRVEIERKSLTOME 209

Qy 208 NRKSLMKIILSE 219  
 Db 210 DVLKALQMKLWE 221

RESULT 6  
 AAG93290  
 ID AAG93290 standard; protein; 233 AA.  
 XX  
 AC AAG93290;  
 XX  
 DT 13-SEP-2001 (first entry)  
 XX  
 DE Human protein HP10650.  
 XX  
 DE Human, gene therapy; tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200142302-A1.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 06-DEC-2000; 2000WO-JP008631.  
 XX  
 PR 06-DEC-1999; 99JP-00346863.  
 PR 06-DEC-1999; 99JP-00346864.  
 PR 08-FEB-2000; 2000JP-00031062.  
 PR 10-FEB-2000; 2000JP-00034090.  
 PR 10-FEB-2000; 2000JP-00034091.  
 PR 14-FEB-2000; 2000JP-00035829.  
 PR 14-FEB-2000; 2000JP-00035839.  
 PR 14-MAR-2000; 2000JP-00071161.  
 PR 30-MAY-2000; 2000JP-00160851.  
 XX  
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 XX  
 PI Kato S, Eguchi C, Saeki M;  
 XX  
 DR WPI, 2001-381646/40.  
 DR N-PSDB; AAH68575.

Human protein originated from tumor cell line, applicable as drug,  
 PT reagent for studying intracellular protein networks and protein source  
 PT for drug screening, also encoded cDNA for gene diagnosis and gene  
 PT therapy.

Claim 1; Page 303-304; 471bp; Japanese.

XX The present sequence is a human protein. The human protein, preferably  
 CC originated from tumor cell line, is applicable as a drug, a reagent for  
 CC studying intracellular protein networks and a protein source for  
 CC screening proteins for binding low molecular weight drugs. The human  
 CC protein coding sequence is useful for gene diagnosis and gene therapy,  
 CC expression vectors and transformant cells for detection of ligands and  
 CC receptors  
 CC  
 XX

[illegible]

	CC	present invention describes a method for identifying such proteins that
	CC	can be used in the treatment of various diseases associated with altered
	CC	mitochondrial function including diabetes mellitus, Huntington's disease,
	CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
	CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
	CC	ragged red fibre syndrome (MERF) or cancer. Accordingly, these
	CC	compositions have neuroprotective, neurotropic, antidiabetic,
	CC	anticonvulsant, antiarrhythmic, osteopathic, ophthalmological and
	CC	cyclostatic activities. This polypeptide sequence is a human heart
	CC	mitochondrial protein of the invention.
Dx		
SQ	Sequence 233 AA:	
Query Match	9.4%; Score 113; DB 7; Length 233;	
Best Local Similarity	27.1%; Pred. No. 0.0011;	
Matches	52; Conservative 30; Mismatches 88; Indels 22; Gaps 8	
Dy		
31	ASPTTMEWDYOVKSSSPFLGAGAEPPAGPOLPSWLPQRCFAVFOCAOCHAVLADS 90	
	: :	
Dz	49 ASMWSSMSNSDASY----ADMERAPQL--EERAAALAE-----ERPVLFLSGGRRLPGDS 95	
Oy	91 VHLAWDLSR-SLGAVFSRVTNNVVLEAPFLVGISLKGSTYNLLFCSCGCIPIVFPHLY 149	
	: :	
Dd	96 --LSWVASGDENNCILLRCVCSNVSVDREOKILSRREKNGCVLETTLCCACCSINLTGVYR 153	
Oy	150 STTAALLAARGHFCLSSDDKVCYL--RTKAIVNASEMDIOWNPVLSKIALKEKIYLVTH 207	
	: :	
Dd	154 CTFKNLDYRRDFCLSLVEAIIESYVLGSSEKOIV-SEDKELPNL--ESRVIEKSILTOME 209	
Oy	208 NRTLKSIIMKLISE 219	
	: :	
Dd	210 DVAKALQMKTWE 221	

RESULT 8	
ADL31517	
ID	ADL31517 standard; protein; 233 AA.
XX	
XX	
AC	ADL31517;
DT	20-MAY-2004 (first entry)
XX	
DE	Human protein encoded by a full length cDNA clone SeqID 3550.
XX	
KW	human; medicine; signal transduction; glycoprotein; transcription; oligo-capping method.
KW	
OS	Homo sapiens.
XX	
PN	EP1396543-A2.
XX	
PD	10-MAR-2004.
XX	
PF	07-JUL-2000; 2003EP-00025638.
XX	
PR	08-JUL-1999; 99JP-00194486.
PR	11-JAN-2000; 2000JP-00118774.
PR	02-MAY-2000; 2000JP-00183865.
PR	07-JUL-2000; 2000EP-00114089.
XX	
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.
XX	
PI	Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y; Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
PI	
DR	WPI, 2004-204755/20.
XX	
DR	N-PSDB; ADL31516.
XX	
PT	New oligonucleotide primers (830 CDNAs) useful for synthesizing full length human cDNAs.
XX	
PS	Example 1; SEQ ID NO 3550; 1340bp; English.
XX	

CC This invention relates to a novel primers useful for synthesising full  
CC length cDNA molecules that encode human proteins. Specifically, it refers  
CC to secretory or membrane proteins that are potential therapeutic agents/  
CC target molecules in the field of medicine, and in particular genes  
CC encoding proteins that are associated with signal transduction,  
CC glycoproteins and transcription. The present invention describes a method  
CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
CC ends using the oligo-capping method. This polypeptide sequence is a full  
CC length human protein of the invention.

XX Sequence 233 AA;

Query Match 9.4%; Score 113; DB 8; Length 233;

Best Local Similarity 27.1%; Pred. No. 0.0011; Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTSMEDIQVYVSSPLGPAGAEPAAPOLPQWLOPERCAVFCQACHAVLADS 90  
DB 49 ASMSMSMEDASV-----ADMERARQ--REEMAAAE-----ERPLVFLSGGCRPLGDS 95  
QY 91 VHLAWDLNR-SIGAVVFSRVTNNVVLVLEAPLVGIEGSLKSTYNNLLFCGSCGIPVGFHLX 149  
DB 96 --LSWVASQEDTNCILRCVSCNVSVDKEOKLSKREKENGCVLETLCCAGCSLNLGYVYR 153  
QY 150 STHALALRGHFCISSDKMVCYL--KTKAIVNASENDIQVPLSEKIAELKEKIVLTH 207  
DB 154 CTPKRLDYKRDLCFLCSVEAIESYVLGSSERQIV--SEDKLEFNL--ESRVEIKSLTOME 209  
QY 208 NRLKSLMKILSE 219  
DB 210 DVLKRLQMKLWE 221

RESULT 9

ABB54167  
ID ABB54167 standard; protein; 695 AA.

XX ABB54167;

XX 29-AUG-2003 (revised)

XX 16-MAY-2002 (first entry)

XX Lactococcus lactis protein yihC.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis; IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-00004630.

XX 11-APR-2000; 2000FR-00004630.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus  
XX lactis and related species.

XX Claim 6; SEQ ID NO 869; 2504bp; French.

XX The present invention is related to a Lactococcus lactis nucleotide  
XX sequence (ABA90521) and related proteins (ABB53500-ABB56211). The nucleic  
XX acid sequence is useful in the detection and/or amplification of nucleic  
XX acid sequence, particularly to identify Lactococcus lactis or related  
XX species. The proteins of the invention are useful for the biosynthesis or  
XX biodegradation of a composition of interest. The invention helps research

CC in lactic bacteria, particularly useful in the production of yogurt and  
CC cheese. Note: The sequence data for this patent is based on equivalent  
CC patent W020017734 (published 18-OCT-2001) which is available in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 29-AUG-2003 to  
CC standardise OS field)

XX Sequence 695 AA;

Query Match 7.7%; Score 92; DB 5; Length 695;

Best Local Similarity 25.3%; Pred. No. 1.5; Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;

QY 101 LGAVVFSRVNNVVLVLEAPLVGIEGSLKSTYNNLLFCGSCGIPVGFHLSHTHALALRG 160  
DB 101 LGITIF-----FYSGTFPFSGAKGELSKRPMMLITNGITVAY-AYSVYATIMSLNG 153  
QY 161 HF-----CLSSDKMVCYLKTKAIVNASR--MDI-----QNVPLSE-K 195  
DB 154 HMGNNFWFELATLIVIMLIGHLEIMKAIMGADALKDLASLVPRKAKHLSKGVLESELK 213  
QY 196 IABLKEKIVLTHNRLKSLMKILSEVTPDQS 225  
DB 214 VGDL--LVVXENKIPADGLISEALVDES 241

RESULT 10

ADS29356  
ID ADS29356 standard; protein; 695 AA.

XX ADS29356;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #18389.

XX Recombinant DNA construct; transformed plant; improved plant property;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
XX pathogen tolerance; pest tolerance; plant disease resistance;  
XX cell cycle pathway modification; plant growth regulator;  
XX homologous recombination; seed oil yield; protein yield; carbohydrate;  
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
XX bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 18389; 122bp; English.  
XX The invention relates to a recombinant DNA construct comprising a  
XX promoter functional in a plant cell, where the promoter is positioned to  
XX provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 695 AA:

Query Match 7.7%; Score 92; DB 8; Length 695;  
Best Local Similarity 25.3%; Pred. No. 1.5;  
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;

QY 101 LGAVVSRYNNVLEAPLVGIRGSLKSTYMLFCGSGCIPVGFHYSTHAAALRG 160  
DB 101 LGITIF-----FYSGTPFSGAKGELSKRPAMMLITGIVAY-AISVYATISLNG 153  
QY 161 HF-----CLSDPMVCYLKTKAIVNASR-MDI-----QNVPLSR-K 195  
DB 154 HMGNNFWEFLATIVILILGHLIEMKAIMGAGALKDIALVPRKAKHKGKQVLESELK 213  
QY 196 IABLKERVLTNRRLKSLMKLILSEVTPDQ 225  
DB 214 VGDL--LLVKNEXKIPADGLITSEALVDES 241

RESULT 11

AA008972 ID AA008972 standard; protein; 269 AA.

XX AC AA008972;

DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 22864.

XX KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KM tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI, 2001-514838/56.

XX DR N-PSDB; AA186903.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
XX and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 22864; 1399BP + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and  
XX the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activity/inhibit activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation. Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic format  
XX directly from WIPO at [fwp.wipo.int/pub/published\\_pct\\_sequences](http://fwp.wipo.int/pub/published_pct_sequences)

XX Sequence 269 AA:

Query Match 7.5%; Score 90; DB 4; Length 269;  
Best Local Similarity 23.0%; Pred. No. 0.58;  
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HSRCATPPRGDFC-----GTERAIDQA-----SFTTSMKMDQVKGSSPL----- 50  
DB 39 YRQAAQPH---CPAGEEPGAPQALGDPASTSVSLTAVDDYVCPLGSHALCTCCFQ 95  
QY 51 ---GPGGLAEERPAAGPOLPSWLQPERCAVFCQACQCHAVLADSVHLAMDLSR----- 99  
DB 96 PMPDRPAERERQDPRVAPQ-----QCAVC---LQPCCHLYMGCTRTGCGCCLA 139  
QY 100 ----SLGAVVSRYNNVLEAPLVGIRGSLKSTY-NLFCGSGCIPVGFHYSTHA 153  
DB 140 PFCENLGDGKCIDGVANNNSYSDDLKNTVLT-RGLTWQML-----TTS 183  
QY 154 ALAALRGHFCLS-----SPRMVCYLKTKAIVNASRMDIQNVPLSR 194  
DB 184 LMALGRGVFLSDRYVTGDTVLCCGLRSFRELTVQYRONIPASE 229

RESULT 12

ABM83354 ID ABM83354 standard; protein; 572 AA.

XX AC ABM83354;

DT 18-NOV-2004 (first entry)

XX DE Human diagnostic and therapeutic protein SEQ ID NO:3603.

XX KM gene therapy; human diagnostic and therapeutic polynucleotide; dthp.

XX OS Homo sapiens.

XX PN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,

XX PI Hartshorne TA, Suchorolski MT, Altus CM, Plets SJ, Elder LV;

XX PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;

XX PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH,

XX PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;

PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Pacury S, Shi X, Suarez CJ;  
PI  
DR WPI; 2004-329368/30.  
DR N-PSDB; ACN42006.  
XX  
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
PS Claim 27; Page; 190pp; English.  
XX  
CC The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dihp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders, endocrine  
CC autoimmune/inflammatory disorder, developmental disorder, or  
CC disorders, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dihp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dihp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 572 AA;  
Query Match 7.4%; Score 89; DB 8; Length 572;  
Best Local Similarity 23.0%; Pred. No. 2.4;  
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;  
QY 8 HRSRCATPPRRDFC-----GCTERAIDQA-----SFTSMEMDPOVYKSSPL----- 50  
DB 342 YRRQAAPPH---CPABEGEPAPQALGDAPSTSVSLTTAVQDYVCPLOGSHALCTCCFQ 398  
QY 51 ---GPAGIGAEPPAAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLR----- 99  
DB 399 PMPDRARERBDPRVAPQ-----QCAVC---LQPFCHLYMGCRTGCGCLA 442  
QY 100 -----SLGAVFSRYTNVNVLEAPFLVIGESLKGSTY-NLLFGSGCGIPVGFHLYSTHA 153  
DB 443 PFCENLNGDKCLDGVLTNNNSYESDILKNYLAT-RGLTWKNML-----TES 486  
QY 154 ALAALRGHFCLS-----SDKMVCYLKTKATVNASMDIONVPLSE 194  
DB 487 LVALLQGVFLSDYRVGTDTLVCYCCGLRSFRELTYOQNIPIASE 532  
RESULT 13  
ID AAB93182 standard; protein; 623 AA.  
XX  
AC AAB93182;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:12128.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-00116126.

XX  
PR 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-0018776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
PT length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.  
XX  
PS Claim 8; SEQ ID NO 12128; 2537pp + Sequence listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination  
CC of the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AA992446 to AA95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention  
XX  
SQ Sequence 623 AA;  
Query Match 7.4%; Score 89; DB 4; Length 623;  
Best Local Similarity 23.0%; Pred. No. 2.7;  
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;  
QY 8 HRSRCATPPRRDFC-----GCTERAIDQA-----SFTSMEMDPOVYKSSPL----- 50  
DB 393 YRRQAAPPH---CPABEGEPAPQALGDAPSTSVSLTTAVQDYVCPLOGSHALCTCCFQ 449  
QY 51 ---GPAGIGAEPPAAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLR----- 99  
DB 450 PMPDRARERBDPRVAPQ-----QCAVC---LQPFCHLYMGCRTGCGCLA 493  
QY 100 -----SLGAVFSRYTNVNVLEAPFLVIGESLKGSTY-NLLFGSGCGIPVGFHLYSTHA 153  
DB 494 PFCENLNGDKCLDGVLTNNNSYESDILKNYLAT-RGLTWKNML-----TES 537  
QY 154 ALAALRGHFCLS-----SDKMVCYLKTKATVNASMDIONVPLSE 194  
DB 538 LVALLQGVFLSDYRVGTDTLVCYCCGLRSFRELTYOQNIPIASE 583  
RESULT 14  
ID ABB97233  
XX  
AC ABB97233 standard; protein; 623 AA.  
XX  
XX



DT 27-JUN-2002 (first entry)  
 XX  
 DE Novel human protein SEQ ID NO: 501.  
 XX  
 KW Human; anti-neoplastic; anti-inflammatory; immunomodulator;  
 KW anti-infective; cerebroprotective; cytoskeletal; rheumatic; gene therapy;  
 KW neuroprotective; anti-Parkinsonian; protein therapy; BSI;  
 XX expressed sequence tag.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200222660-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 10-SEP-2001; 2001MO-US026015.  
 XX  
 PR 11-SEP-2000; 2000US-00659671.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Aeundi V, Zhang J, Zhao QA, Ren F,  
 PI xue AJ, Yang Y, Wehrman T, Dermanac RT;  
 DR MPI; 2002-292408/33.  
 DR N-PSDB; ABN32419.  
 XX  
 PT An isolated polynucleotide for treating diseases associated with its  
 PT encoded polypeptide such as cancer and multiple sclerosis.  
 XX  
 PS Example 2; SEQ ID NO 501; 509bp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of 444  
 CC novel human proteins. These were isolated from expressed sequence tags  
 CC (ESTs). They can be used to stimulate cell growth, to regulate  
 CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
 CC Parkinson's disease. The present sequence is a protein of the invention  
 XX  
 XX Sequence 623 AA;  
 SQ  
 Query Match 7.4%; Score 89; DB 5; Length 623;  
 Best Local Similarity 23.0%; Pred. No. 2.7;  
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;  
 QY 8 HRSRCATPPRGDFC-----GTERAIDQA-----SFTSMEMDTQVKGSSPL----- 50  
 DB 393 YRRQAQAPRH---CPAPEGEFGAPQALGDAPRTSVSLTTAVQDVCPQLQSHALCTCCFQ 449  
 QY 51 ---GPAGLGAEPAPQPSWLOPERCAVFOCAQCAVLAADSHLAMDLSR----- 99  
 DB 450 PMPDRRAEREDPRVAPO-----QCAVC---LQPFCHLYMGCTRGTCYGCCLA 493  
 QY 100 -----SIGAVVFSRVTNVNLVLEAPFLVIGSLKSTY-NLFCGSGGIPVGFHLVSTHA 153  
 DB 494 PFCFLNIGDKCLDGLVANNNSYESDILKNYLAT-RGLTWKKNML-----TES 537  
 QY 154 ALAALRGHFLS-----SDKNVCYLTKTKALVNASSEMDIONVPLSE 194  
 DB 538 LVALQRGVFLSDYRVYGVDTVLCTCCGGRSFRRLTYQYQNIIPASE 583  
 RESULT 15  
 ID AAB93168  
 AC AAB93168; protein; 652 AA.  
 XX  
 DT 26-JUN-2001 (first entry)

XX  
 DE Human protein sequence SEQ ID NO:12100.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-00116126.  
 XX  
 PR 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR MPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX  
 PS Claim 8; SEQ ID NO 12100; 2537pp + Sequence listing; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH11633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 XX  
 XX Sequence 652 AA;  
 SQ  
 Query Match 7.4%; Score 89; DB 4; Length 652;  
 Best Local Similarity 23.0%; Pred. No. 2.9;  
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;  
 QY 8 HRSRCATPPRGDFC-----GTERAIDQA-----SFTSMEMDTQVKGSSPL----- 50  
 DB 422 YRRQAQAPRH---CPAPEGEFGAPQALGDAPRTSVSLTTAVQDVCPQLQSHALCTCCFQ 478  
 QY 51 ---GPAGLGAEPAPQPSWLOPERCAVFOCAQCAVLAADSHLAMDLSR----- 99  
 DB 479 PMPDRRAEREDPRVAPO-----QCAVC---LQPFCHLYMGCTRGTCYGCCLA 522  
 QY 100 -----SIGAVVFSRVTNVNLVLEAPFLVIGSLKSTY-NLFCGSGGIPVGFHLVSTHA 153  
 DB 523 PFCFLNIGDKCLDGLVANNNSYESDILKNYLAT-RGLTWKKNML-----TES 566

Mon Nov 21 12:44:37 2005

**us-09-942-052a-728.rag**

Page 10

```

Oy      154 ALAALRGHFCIS-----SDKWCYLLKTKAIVNASENDIQNPPLSE 194
          :| | | | | :| | | :| | |
Db      567 LVALQGVFLTSDYRVGTIVLCYCCGHSFRELTYQYRQNIPIASE 612

```

Search completed: November 18, 2005, 20:31:06  
Job time : 196 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 20:24:56 / Search time 48 Seconds  
(without alignments)  
394,432 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAOPLRHRSRCATPPRGDF.....LKSIMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5.COMB.dep:\*
- 2: /cgn2\_6/prodata/1/iaa/6.COMB.dep:\*
- 3: /cgn2\_6/prodata/1/iaa/H.COMB.dep:\*
- 4: /cgn2\_6/prodata/1/iaa/PCITUS.COMB.dep:\*
- 5: /cgn2\_6/prodata/1/iaa/RE.COMB.dep:\*
- 6: /cgn2\_6/prodata/1/iaa/backfile1.dep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	7.2	158	US-09-134-000C-6596	Sequence 6596, Ap
2	83	6.9	457	5268463-7	Patent No. 5268463
3	81.5	6.8	456	5432081-7	Patent No. 5432081
4	81.5	6.8	457	US-08-882-704A-6	Sequence 6, Appl1
5	81.5	6.8	457	US-09-151-957-6	Sequence 6, Appl1
6	80.5	6.7	1711	US-08-342-930-2	Sequence 3, Appl1
7	78.5	6.6	325	US-09-018-576-3	Sequence 3, Appl1
8	78.5	6.6	325	US-09-018-576-12	Sequence 12, Appl1
9	78.5	6.6	325	US-09-248-137-3	Sequence 12, Appl1
10	78.5	6.6	325	US-09-248-137-12	Sequence 12, Appl1
11	78.5	6.6	452	US-09-771-161A-237	Sequence 237, Appl1
12	78	6.5	190	US-09-252-991A-19049	Sequence 19049, A
13	77	6.4	607	US-09-537-682-1	Sequence 1, Appl1
14	77	6.4	719	US-08-520-933-3	Sequence 3, Appl1
15	77	6.4	719	US-09-285-040-3	Sequence 3, Appl1
16	77	6.4	738	5262177-2	Patent No. 5262177
17	77	6.4	776	US-09-949-016-10829	Sequence 10829, A
18	76.5	6.4	282	US-09-845-713A-2	Sequence 2, Appl1
19	76	6.3	351	US-09-902-540-13214	Sequence 13214, A
20	76	6.3	1001	US-10-104-047-2283	Sequence 2283, Ap
21	75.5	6.3	310	US-08-129-456A-36	Sequence 36, Appl1
22	75.5	6.3	310	US-08-705-868-3	Sequence 3, Appl1
23	75.5	6.3	310	US-09-123-615-3	Sequence 3, Appl1
24	75.5	6.3	310	US-08-360-821B-35	Sequence 35, Appl1
25	75.5	6.3	310	US-09-851-026-35	Sequence 35, Appl1
26	75.5	6.3	1449	US-08-840-062-6	Sequence 6, Appl1
27	75.5	6.3	1723	US-09-194-612A-31	Sequence 31, Appl1

28	75.5	6.3	4545	1	US-08-804-227C-14	Sequence 14, Appl1
29	75.5	6.3	4550	1	US-08-804-227C-8	Sequence 8, Appl1
30	75.5	6.3	4550	1	US-08-804-198-2	Sequence 2, Appl1
31	75	6.3	78	2	US-09-270-767-43496	Sequence 43496, A
32	75	6.3	458	2	US-09-800-170-3	Sequence 3, Appl1
33	75	6.3	717	6	5262177-5	Patent No. 5262177
34	75	6.3	784	2	US-09-724-586A-1	Sequence 1, Appl1
35	75	6.3	784	2	US-09-600-823-1	Sequence 1, Appl1
36	75	6.3	784	2	US-09-724-666-1	Sequence 1, Appl1
37	75	6.3	1088	2	US-09-920-804-2	Sequence 2, Appl1
38	75	6.3	1088	2	US-09-920-804-4	Sequence 2, Appl1
39	75	6.3	1129	2	US-09-734-674-2	Sequence 2, Appl1
40	75	6.3	1129	2	US-10-274-990-2	Sequence 2, Appl1
41	75	6.3	2227	2	US-08-475-886-4	Sequence 4, Appl1
42	75	6.3	2227	2	US-09-653-499-4	Sequence 4, Appl1
43	74.5	6.2	211	2	US-10-135-988-4	Sequence 4, Appl1
44	74.5	6.2	211	2	US-09-949-016-11409	Sequence 11409, A
45	74.5	6.2	279	2	US-09-107-433-3335	Sequence 3335, Ap

## ALIGNMENTS

```
RESULT 1
US-09-134-000C-6596
Sequence 6596, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6596
LENGTH: 158
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-6596

Query Match      7.2%; Score 86; DB 2; Length 158;
Best Local Similarity 26.5%; Pred. No. 0.18;
Matches 40; Conservative 27; Mismatches 54; Indels 30; Gaps 9;

QY      87 LADSVHLANDLSRLGAVY---FSRYTNVVEA--PLVIGISLKG-----STYN 133
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      17 LREYIYLAEEQPLSTGISTLNGLSKPTGNLNAFFPFINDEGDPKGFSTSNYIKTAN 76
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      134 LIFCGSGGIPVGFHLVSTYAAALAL---RDHCLSSDKWVCYLTKTKAVNASEMDIQNV 190
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      77 YQVCSGCSIPSDNQNETSPSAIALLGYIRNF-----KNTC---TKVLTLALNGFENK 128
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      191 PLSF---KIAIEKXIVLTHNRKLSMKIL 217
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      129 PISKVTRKLSIDLSYK-DLYQEMK-LTKII 157
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
5268463-7
Patent No. 5268463
APPLICANT: JEFFERSON, RICHARD A.
TITLE OF INVENTION: PLANT PROMOTER a-GLUCURONIDASE GENE
CONSTRUCT
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/447,976
FILING DATE: 08-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 119,102
```

```
; FILING DATE: 10-NOV-1987
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; SEQ ID NO: 7
; LENGTH: 457
5268463-7

Query Match      6.8%; Score 83; DB 6; Length 457;
Best Local Similarity 25.2%; Pred. No. 2;
Matches 37; Conservative 30; Mismatches 54; Indels 26; Gaps 8;

Oy 95 WDLSSLAGAVFSRVTNNV-----VLEA-----PVLVIGESLKGSTYNLL-FCGSCG 141
Db 316 WSLPVALVALAIASIGQVMTVMMALEADTVYEGEYLTVR--IEGLTVSLFSTRKCG 373
Oy 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLKTKKAIYNASEMDIQNV-----PLSE 194
Db 374 QAIGGS--SIPAFILGLSGYIANOVQTPREVIMGIRTSIALVPCGFMALAFYIWFYPLTD 431
Oy 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KFEKIVEIDNRKRVQOQLISDIT 456

RESULT 3
5432081-7
; Patent No. 5432081
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: HOST CELLS TRANSFORMED WITH THE E. COLI
; GLUCURONIDE PERMEASE GENE
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,546
; FILING DATE: 15-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 447,976
; FILING DATE: 08-DEC-1989
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; APPLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; SEQ ID NO: 7
; LENGTH: 456
5432081-7

Query Match      6.8%; Score 81.5; DB 6; Length 456;
Best Local Similarity 25.2%; Pred. No. 2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

Oy 95 WDLSSLAGAVFSRVTNNV-----VLEA-----PVLVIGESLKGSTYNLL-FCGSCG 141
Db 316 WSLPVALVALAIASIGQVMTVMMALEADTVYEGEYLTVR--IEGLTVSLFSTRKCG 373
Oy 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLKTKKAIYNASEMDIQNV-----PLSE 194
Db 374 QAIGGS--SIPAFILGLSGYIANOVQTPREVIMGIRTSIALVPCGFMALAFYIWFYPLTD 430
Oy 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 431 K--KFEKIVEIDNRKRVQOQLISDIT 455

RESULT 4
US-08-882-704A-6
; Sequence 6, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: JEFFERSON, RICHARD A.
; APPLICANT: WILSON, KATHERINE J.
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,704A
FILING DATE: 25-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5879906tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4800
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-882-704A-6

Query Match      6.8%; Score 81.5; DB 1; Length 457;
Best Local Similarity 25.2%; Pred. No. 2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

Oy 95 WDLSSLAGAVFSRVTNNV-----VLEA-----PVLVIGESLKGSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIASIGQVMTVMMALEADTVYEGEYLTVR--IEGLTVSLFSTRKCG 374
Oy 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLKTKKAIYNASEMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGLSGYIANOVQTPREVIMGIRTSIALVPCGFMALAFYIWFYPLTD 431
Oy 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KFEKIVEIDNRKRVQOQLISDIT 456

RESULT 5
US-09-151-957-6
; Sequence 6, Application US/09151957
; Patent No. 6429292
; GENERAL INFORMATION:
; APPLICANT: JEFFERSON, RICHARD A.
; APPLICANT: WILSON, KATHERINE J.
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,957
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
```

```

/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/882,704
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 6429292cendburg Ph.D., Carol
/ REGISTRATION NUMBER: 39,317
/ REFERENCE/DOCKET NUMBER: 190106.404
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-6900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 457 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-151-957-6

Query Match 6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred.No.2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8

Cy 95 WDLSSLGAVVSRYTNV-----VLEA-----PVLVIGESLSGSTYNLL-FCGSCG 141
Db 317 WSLPVALLAALAIASIGQVTTVMALADYVEYGLVGR--IGLITSLSFTRKCG 374
Cy 142 IPVGFHLYSTHAALALRGHFC--LSSDMVCYLLTKKAIYVASENDIONV-----PLSE 194
Db 375 QAIG---GSIPAFILGSGYINQVOTPEVINGIRTSIALVPCGFLLAVIIWPFPLTD 431
Cy 195 KIAELKEKIVLTHNRKLSIMKILSEVT 221
Db 432 K--KEKEIIVEIDNRKVOQOOLSDIT 456

RESULT 6
US-08-342-930-2
/ Sequence 2, Application US/08342930
/ Patent No. 5821084
/ GENERAL INFORMATION:
/ APPLICANT: OLNSTED, ELIZABETH A.
/ APPLICANT: MAURO, LAURA J.
/ APPLICANT: DAVIS, ALAN R.
/ APPLICANT: DIXON, JACK E.
/ TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE
/ TITLE OF INVENTION: PHOSPHATASE
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORRISON & FOERSTER
/ STREET: 755 Page Mill Road
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304-1018
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/342,930
/ FILING DATE: 21-NOV-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KONSKI, ANTOINETTE F.
/ REGISTRATION NUMBER: 34,202
/ REFERENCE/DOCKET NUMBER: 20344-20975.00
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 813-5600
/ TELEFAX: (415) 494-0792
/ TELER: 706141
/ INFORMATION FOR SEQ ID NO: 2:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-342-930-2

Query Match          6.7%; Score 80.5; DB 1; Length 1711;
Best Local Similarity 29.6%; Pred. No. 28;
Matches 21; Conservative 8; Mismatches 17; Indels 25; Gaps 1;

QY      36 SMEWDTVVKGSSPLGAGLGAEEBPAAQPOLPSWLQPERCAVFQOAOCHAVLADSVHLAW 95
       ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      753 SMQGSASVNLWMPSSGLGGQA-----CHAQLSDAGHLSW 787

QY      96 DLSRSLGAVVF 106
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      788 EQPKLGQELF 798

RESULT 7
US-09-018-576-3
Sequence 3, Application US/09018576
Patent No. 5968800
GENERAL INFORMATION:
APPLICANT: Gerhold, David L.
TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, RY60-30
CITY: Rahway
STATE: NJ
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/018,576
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Hand, J. Mark
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 19885Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732/594-3905
TELEFAX: 732/594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-018-576-3

Query Match          6.6%; Score 78.5; DB 1; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.8;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;

QY      22 GATERAIQOAFETTSMEWDTVVVGSSPLGPAGLG-----AEEPAG 63
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      43 GPNDALAEITKALGEMD-DNQYVVQAKVFPHGGGFVLAEFFMLSDLAEVVRHQRPLAQ 101

QY      64 POLPSMWLOPERCAVFQOAOCHAVLADSVHLAMDLSRLSGAVVFSRVTNNVF---LEAPF 119
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      102 AQVSXYLQMLKGV---AFCHA-----NNIVHRDLKPANL 133

120 LVGIKSLKSGTYNLLPFGSCGIPVGRHTLYSTHAALALRGHFCLSSD----- 167

```

Db 134 LISASGQKIAIDFGIARVFS---PDGSRLY-THQ--VATRSVGCIMGELINGSPLEPGKN 187  
QY 168 --KAVCYLLKTKAIYNASEM-DIONVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQ 224  
Db 188 DIEQLCYVLRILGTNPQVWPELTLPDYNKIS-FKEQVPMV-----LEEVLDPDV 236  
QY 225 S 225  
Db 237 S 237

RESULT 8  
US-09-018-576-12  
Sequence 12, Application US/09018576  
Patent No. 5968800  
GENERAL INFORMATION:  
APPLICANT: Gerhold, David L.  
TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, RY60-30  
CITY: Rahway  
STATE: NJ  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018,576  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hand, J. Mark  
REGISTRATION NUMBER: 36,545  
REFERENCE/DOCKET NUMBER: 19885Y  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732/594-3905  
TELEFAX: 732/594-4720  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-018-576-12

Query Match 6.6%; Score 78.5; DB 1; Length 325;  
Best Local Similarity 23.2%; Pred. No. 3.8;  
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;

QY 22 GGTARAIOASFTTSMEMDQVYVKGSSPLGPAAG-----AEEPAAG 63  
Db 43 GEPNOLAREIKALOEME-DNQYVVLKAVFPHGGFVLAFEMISDLAEVVRHAPRLAQ 101  
QY 64 POLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVVFSRVTVNVV-----LEAPF 119  
Db 102 AQVKSIVLQMLKGV---AFCHA-----NNIVHRDLKPA NL 133  
QY 120 LVGIGSLKSGSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFCLSDD----- 167  
Db 134 LISASGQKIAIDFGIARVFS---PDGSRLY-THQ--VATRSVGCIMGELINGSPLEPGKN 187  
QY 168 --KAVCYLLKTKAIYNASEM-DIONVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQ 224  
Db 188 DIEQLCYVLRILGTNPQVWPELTLPDYNKIS-FKEQVPMV-----LEEVLDPDV 236  
QY 225 S 225

Db 237 S 237

RESULT 9  
US-09-248-137-3  
Sequence 3, Application US/09248137  
Patent No. 6030788  
GENERAL INFORMATION:  
APPLICANT: Gerhold, David L.  
TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, RY60-30  
CITY: Rahway  
STATE: NJ  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/248,137  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/018,576  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hand, J. Mark  
REGISTRATION NUMBER: 36,545  
REFERENCE/DOCKET NUMBER: 19885Y  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732/594-3905  
TELEFAX: 732/594-4720  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-248-137-3

Query Match 6.6%; Score 78.5; DB 2; Length 325;  
Best Local Similarity 23.2%; Pred. No. 3.8;  
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;

QY 22 GGTARAIOASFTTSMEMDQVYVKGSSPLGPAAG-----AEEPAAG 63  
Db 43 GEPNOLAREIKALOEME-DNQYVVLKAVFPHGGFVLAFEMISDLAEVVRHAPRLAQ 101  
QY 64 POLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVVFSRVTVNVV-----LEAPF 119  
Db 102 AQVKSIVLQMLKGV---AFCHA-----NNIVHRDLKPA NL 133  
QY 120 LVGIGSLKSGSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFCLSDD----- 167  
Db 134 LISASGQKIAIDFGIARVFS---PDGSRLY-THQ--VATRSVGCIMGELINGSPLEPGKN 187  
QY 168 --KAVCYLLKTKAIYNASEM-DIONVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQ 224  
Db 188 DIEQLCYVLRILGTNPQVWPELTLPDYNKIS-FKEQVPMV-----LEEVLDPDV 236  
QY 225 S 225  
Db 237 S 237

RESULT 10

US-09-248-137-12  
 ; Sequence 12, Application US/09248137  
 ; Patent No. 6030788  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gerhold, David L.  
 ; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merck & Co., Inc.  
 ; STREET: P. O. Box 2000, RY60-30  
 ; CITY: Rahway  
 ; STATE: NJ  
 ; COUNTRY: US  
 ; ZIP: 07065-0907  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/248,137  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/018,576  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hand, J. Mark  
 ; REGISTRATION NUMBER: 36,545  
 ; REFERENCE/DOCKET NUMBER: 19885Y  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 732/594-3905  
 ; TELEFAX: 732/594-4720  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 325 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-248-137-12

Query Match 6.6%; Score 78.5; DB 2; Length 325;  
 Best Local Similarity 23.2%; Pred. No. 3.8;  
 Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;  
 QY 22 GTERAIDQASFTSMEMDTQVVGSSPLGPGAGL-----AEEPAG 63  
 DB 43 GEPNQLREIKALQEME-DNQYVQLKAVFPHGGFVLAFEFMLSDLAEVVRHQRPLAQ 101  
 QY 64 POLPSMLQPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNV-----LEAPF 119  
 DB 102 AOVKSYLQMLKGV---AFCHA-----NNIVHRDLKPANL 133  
 QY 120 LVGIEGSLKSTYMLFCGSGCIPVGFHLVSTHAALALRGHFCSSD----- 167  
 DB 134 LISASGQKIAIDFGLARVFS---PDGSRLY-THQ-VATRSVGTIMGELANGSPFLPGKN 187  
 QY 168 --KAVCYLTKTAIVNASM-DIONVPLSEKIAELKEKIVLTHNRLKSLMKLSEVTPDQ 224  
 DB 188 DIEQLCYVRLIGTPNPQVWPELTLPDYNKIS-FKQVPMF-----LEEVLDPV 236  
 QY 225 S 225  
 DB 237 S 237

RESULT 11  
 US-09-771-161A-237  
 ; Sequence 237, Application US/09771161A  
 ; Patent No. 6936450  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LEVINE, et al.

TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
 ; FILE REFERENCE: 802620-2005.1  
 ; CURRENT APPLICATION NUMBER: US/09/771,161A  
 ; CURRENT FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: 09/724,676  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 136776  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: 135619  
 ; PRIOR FILING DATE: 2000-04-12  
 ; NUMBER OF SEQ ID NOS: 273  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 237  
 ; LENGTH: 452  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-771-161A-237

Query Match 6.6%; Score 78.5; DB 2; Length 452;  
 Best Local Similarity 23.2%; Pred. No. 6.3;  
 Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;  
 QY 22 GTERAIDQASFTSMEMDTQVVGSSPLGPGAGL-----AEEPAG 63  
 DB 43 GEPNQLREIKALQEME-DNQYVQLKAVFPHGGFVLAFEFMLSDLAEVVRHQRPLAQ 101  
 QY 64 POLPSMLQPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNV-----LEAPF 119  
 DB 102 AOVKSYLQMLKGV---AFCHA-----NNIVHRDLKPANL 133  
 QY 120 LVGIEGSLKSTYMLFCGSGCIPVGFHLVSTHAALALRGHFCSSD----- 167  
 DB 134 LISASGQKIAIDFGLARVFS---PDGSRLY-THQ-VATRSVGTIMGELANGSPFLPGKN 187  
 QY 168 --KAVCYLTKTAIVNASM-DIONVPLSEKIAELKEKIVLTHNRLKSLMKLSEVTPDQ 224  
 DB 188 DIEQLCYVRLIGTPNPQVWPELTLPDYNKIS-FKQVPMF-----LEEVLDPV 236  
 QY 225 S 225  
 DB 237 S 237

RESULT 12  
 US-09-252-991A-19049  
 ; Sequence 19049, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 19049  
 ; LENGTH: 190  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-19049

Query Match 6.5%; Score 78; DB 2; Length 190;  
 Best Local Similarity 32.9%; Pred. No. 1.9;  
 Matches 24; Conservative 6; Mismatches 29; Indels 14; Gaps 3;  
 QY 4 QPLRRSRCATPPRRDPCGTERAIDQASFTSMEMDTQVVGSSPLGPGAGLEPPAG 63  
 DB 94 RPCRTRRCVAPRAGGYG---RAGDRAS-----VARSTGFLPAAPGRPRGHC- 139

QY 64 POLPSWLOPERCA 76  
Db 140 PAPERPRRRRCA 152

## RESULT 13

US-09-537-682-1  
Sequence 1, Application US/09537682  
Patent No. 6303357  
GENERAL INFORMATION:  
APPLICANT: TAKEUCHI, Kenichi  
APPLICANT: KOIDE, Yoshinao  
APPLICANT: NAKANISHI, Yuji  
APPLICANT: SUZUKI, Satoru  
TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE, RECOMBINANT  
TITLE OF INVENTION: DNA, AND METHOD FOR PRODUCING MODIFIED  
FILE REFERENCE: A20-121814C/KI  
CURRENT FILING DATE: 2000-03-29  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 607  
TYPE: PRT  
ORGANISM: Enterococcus faecium No. 6303357 7044  
US-09-537-682-1

Query Match 6.4%; Score 77; DB 2; Length 607;  
Best Local Similarity 23.2%; Pred. No. 15;  
Matches 49; Conservative 30; Mismatches 76; Indels 56; Gaps 10;

QY 19 DPGCGTERAIDQASFTTSM-----EWDQVVKGSPLGP---AGLGAEPAPGP 65  
Db 51 DFBEGT-----SRSSTKLHGGIRYKTFDVEVY--ADTGBRAVVGTAIPHIPKDDP 102  
QY 66 LPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVVESRVTN-----NVLEAP 119  
Db 103 LPLIYEBEGATTTFMF-----SVKAVAMDLYDKLANVGTKEVNTLTPBEVLBEPEP 154  
QY 120 LVGIEGSLKSTYNLFCGSCGIFVGF-----HLYSTHAALALRGHFCLSDDMCYLL 174  
Db 155 LK--KEGLKGA-----GYLLDPRNDARLVINDIKKAABDGAIVLVSKKAVGFLY 202  
QY 175 KTKAIVNASEMDIQNVPLSKIAELKEKIYL 205  
Db 203 EGDQIVGVKARDL---LTDEVTIEIKSLVI 229

## RESULT 14

US-08-520-933-3  
Sequence 3, Application US/08520933  
Patent No. 5981194  
GENERAL INFORMATION:  
APPLICANT: Jefferies, Wilfred A.  
APPLICANT: McGeer, Patrick L.  
APPLICANT: Rothenberger, Sylvia  
APPLICANT: Food, Michael R.  
APPLICANT: Yamada, Tatsuo  
APPLICANT: Kennard, Malcolm  
TITLE OF INVENTION: Use of p97 and Iron Binding Proteins  
TITLE OF INVENTION: as Diagnostic and Therapeutic Agents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bereskin & Parr  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/520,933  
FILING DATE: August 31, 1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Shona S. McDiarmid  
REGISTRATION NUMBER: 38,798  
REFERENCE/DOCKET NUMBER: 7685-006  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 416-364-7311  
TELEFAX: 416-361-1398  
TELEX: 06-23115  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-520-933-3

Query Match 6.4%; Score 77; DB 1; Length 719;  
Best Local Similarity 23.7%; Pred. No. 19;  
Matches 59; Conservative 25; Mismatches 87; Indels 78; Gaps 15;

QY 17 RDPFGCGTERAIDQASFTTSM-----DTQVVKGSPLGPAGLGAEPAPGP 64  
Db 171 RQDSSG--EGVCDKSLERYDYISGAFRLAGAGDAVAFVKISTYL-----ENTDK 220  
QY 65 QLPFWLP-----BRCAVFOCAQCH--AVLADSVHLAMDLSRSLGAVVESRV 109  
Db 221 TLPFWQALLSQDFELLCRDSRADVTBWRQCHLARVPAAVAVRADTD---GGLIF-RL 276  
QY 110 TNNVLEAPFLVIGRS---LKGSTY---NLLFCGSCG--IPVGHLYST---HAALA 157  
Db 277 LV-----EGQRLSHSGSSPQMFSSRAYGQKLLFKDSTSELVPIATQTYEAMVIGHEYLHA 332  
QY 158 LRGHFCLSDKWCYLL---LTKAIVNASEMDIQNVPLSKIAELKEKIYLTNNRLKSL 213  
Db 333 MKGLIC-DPNRLPIYRMCVLSTPEIQKGDV-----AAVRRQRKKE 375  
QY 214 MKILSEVTP 222  
Db 376 IQCVSAKSP 384

## RESULT 15

US-09-285-040-3  
Sequence 3, Application US/09285040  
Patent No. 6455494  
GENERAL INFORMATION:  
APPLICANT: Jefferies, Wilfred A.  
APPLICANT: McGeer, Patrick L.  
APPLICANT: Rothenberger, Sylvia  
APPLICANT: Food, Michael R.  
APPLICANT: Yamada, Tatsuo  
APPLICANT: Kennard, Malcolm  
TITLE OF INVENTION: Use of p97 and Iron Binding Proteins  
TITLE OF INVENTION: as Diagnostic and Therapeutic Agents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bereskin & Parr  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/285,040  
FILING DATE: 01-APR-1999  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gravelle, Michelle  
REGISTRATION NUMBER: 40,261  
REFERENCE/DOCKET NUMBER: 7685-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 416-364-7311  
TELEFAX: 416-361-1398  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-285-040-3

Query Match 6.4%; Score 77; DB 2; Length 719;

Best Local Similarity 23.7%; Pred. No. 19;  
Matches 59; Conservative 25; Mismatches 87; Indels 78; Gaps 15;

```
QY 17 RQDFCGGTERAIDQASFTSMEW-----DTQVVKSSPLGPGAGAEPPAAGP 64
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 171 RGDSSG--EGVCDKSPLERYDYDYGAFRCLEAGADVAFFVKGSTVL-----ENTDGK 220
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 65 QLPFMLQP-----ERCAVFOCAQCH--AVLADSVHLAMDLSRLGAVFGRV 109
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 221 TLPBWGQALLSQDFELLCRDGSRADVTWRQCHLARVPAAHVVVRADTD---GGILIF-RL 276
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 110 TNNVVLKAPFLVGIKGS---LKGSTY--NILFCGSCG--IPVGFHLYST--HAALAA 157
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 277 LN---EGQRLEFSGSSSFQWFSSEAYGKDLFPDSTSELVPATQTYEAMLGHEYLHA 332
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 158 LRGHFCLSSDKKVCYL---LKTKAIVNASEMDIONVPLSEKIAELKEKIVLTNRRLKSL 213
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 333 MKGILLC-DPNRLPPYLKWCVLSTPEIQKCGDM-----AVAFRRQRLKPE 375
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 214 MKILSEVTP 222
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 376 IQCVSAKSP 384
```

Search completed: November 18, 2005, 20:35:56  
Job time : 50 secs

This Page Blank (unused)



```
RESULT 2
US-09-942-052-729
; Sequence 729, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Mangmao
; APPLICANT: Challita-Bid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85PIB3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 729
; LENGTH: 229
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85PIB3/OIPS
US-09-942-052-729

Query Match          100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLRRSRCAATPPRGDFCGGTERAIDQASFTTSMEMDTQVYKSSPLGAGAE 60
   |||||
DB 1 MAAPLRRSRCAATPPRGDFCGGTERAIDQASFTTSMEMDTQVYKSSPLGAGAE 60

QY 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSIGAVFSRVNNVLEA 120
   |||||
DB 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSIGAVFSRVNNVLEA 120

QY 121 VGIGSLKGSTYNNLLFCGSCGIPVGPHLYSTHAALALRGHFCCLSDMVCYLKTK 180
   |||||
DB 121 VGIGSLKGSTYNNLLFCGSCGIPVGPHLYSTHAALALRGHFCCLSDMVCYLKTK 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDQSKPEN 229
   |||||
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDQSKPEN 229

RESULT 3
US-09-942-052-730
; Sequence 730, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Mangmao
; APPLICANT: Challita-Bid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85PIB3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 730
; LENGTH: 229
; TYPE: PRT
```

```
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85PIB3 protein
; OTHER INFORMATION: sequence
US-09-942-052-730

Query Match          100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLRRSRCAATPPRGDFCGGTERAIDQASFTTSMEMDTQVYKSSPLGAGAE 60
   |||||
DB 1 MAAPLRRSRCAATPPRGDFCGGTERAIDQASFTTSMEMDTQVYKSSPLGAGAE 60

QY 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSIGAVFSRVNNVLEA 120
   |||||
DB 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSIGAVFSRVNNVLEA 120

QY 121 VGIGSLKGSTYNNLLFCGSCGIPVGPHLYSTHAALALRGHFCCLSDMVCYLKTK 180
   |||||
DB 121 VGIGSLKGSTYNNLLFCGSCGIPVGPHLYSTHAALALRGHFCCLSDMVCYLKTK 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDQSKPEN 229
   |||||
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDQSKPEN 229

RESULT 4
US-09-942-052-731
; Sequence 731, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Mangmao
; APPLICANT: Challita-Bid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85PIB3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 731
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-942-052-731

Query Match          100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLRRSRCAATPPRGDFCGGTERAIDQASFTTSMEMDTQVYKSSPLGAGAE 60
   |||||
DB 1 MAAPLRRSRCAATPPRGDFCGGTERAIDQASFTTSMEMDTQVYKSSPLGAGAE 60

QY 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSIGAVFSRVNNVLEA 120
   |||||
DB 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSIGAVFSRVNNVLEA 120

QY 121 VGIGSLKGSTYNNLLFCGSCGIPVGPHLYSTHAALALRGHFCCLSDMVCYLKTK 180
   |||||
DB 121 VGIGSLKGSTYNNLLFCGSCGIPVGPHLYSTHAALALRGHFCCLSDMVCYLKTK 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDQSKPEN 229
   |||||
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDQSKPEN 229
```

## RESULT 5

US-09-942-052-707

Sequence 707, Application US/09942052

Publication No. US20030170626A1

GENERAL INFORMATION:

APPLICANT: Raitano, Arthur B.

APPLICANT: Paris, Mary

APPLICANT: Hubert, Rene S.

APPLICANT: Afari, Daniel

APPLICANT: Ge, Wangmao

APPLICANT: Challita-Bid, Pia M.

TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3

TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER

FILE REFERENCE: 51158-20028.00

CURRENT APPLICATION NUMBER: US/09/942,052

CURRENT FILING DATE: 2001-08-28

PRIOR APPLICATION NUMBER: 60/228,432

PRIOR FILING DATE: 2000-08-28

NUMBER OF SEQ ID NOS: 744

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 707

LENGTH: 164

TYPE: PRT

ORGANISM: Unknown Organism

FEATURE: OTHER INFORMATION: Description of Unknown Organism: Splice variant 1

OTHER INFORMATION: open reading frame 3 peptide sequence

FEATURE:

NAME/KEY: MOD RES

LOCATION: (44)

OTHER INFORMATION: Unknown amino acid or stop

FEATURE:

NAME/KEY: MOD RES

LOCATION: (50) (51)

OTHER INFORMATION: Unknown amino acid or stop

FEATURE:

NAME/KEY: MOD RES

LOCATION: (67)

OTHER INFORMATION: Unknown amino acid or stop

FEATURE:

NAME/KEY: MOD RES

LOCATION: (93)

OTHER INFORMATION: Unknown amino acid or stop

US-09-942-052-707

Query Match 11.6%; Score 138.5; DB 3; Length 164;

Best Local Similarity 82.1%; Pred. No. 1.2e-05;

Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Db 162 FCISDKNVCYLTKTKAIVNASEMDIONVPLSEKIAELK 200

2 FFLSS-----YLLKTKAIVNASEMDIONVPLSEKIAEVK 35

## RESULT 6

US-09-942-052-704

Sequence 704, Application US/09942052

Publication No. US20030170626A1

GENERAL INFORMATION:

APPLICANT: Raitano, Arthur B.

APPLICANT: Paris, Mary

APPLICANT: Hubert, Rene S.

APPLICANT: Afari, Daniel

APPLICANT: Ge, Wangmao

APPLICANT: Challita-Bid, Pia M.

TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3

TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER

FILE REFERENCE: 51158-20028.00

CURRENT APPLICATION NUMBER: US/09/942,052

CURRENT FILING DATE: 2001-08-28

PRIOR APPLICATION NUMBER: 60/228,432

PRIOR FILING DATE: 2000-08-28

NUMBER OF SEQ ID NOS: 744

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 704

LENGTH: 29

TYPE: PRT

ORGANISM: Unknown Organism

FEATURE: OTHER INFORMATION: Description of Unknown Organism: 85P1B3 peptide

OTHER INFORMATION: sequence

US-09-942-052-704

Query Match 11.5%; Score 138; DB 3; Length 29;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 172 YLLKTKAIVNASEMDIONVPLSEKIAELK 200

1 YLLKTKAIVNASEMDIONVPLSEKIAELK 29

## RESULT 7

US-09-890-688-82

Sequence 82, Application US/09890688

Publication No. US20030144475A1

GENERAL INFORMATION:

APPLICANT: Seishi KATO

APPLICANT: Chikashi EGUCHI

APPLICANT: Mihoro SAeki

TITLE OF INVENTION: Human Proteins and cDNAs thereof

FILE REFERENCE: 2001-11024/WMC/00653

CURRENT APPLICATION NUMBER: US/09/890,688

CURRENT FILING DATE: 2001-08-06

PRIOR APPLICATION NUMBER: JP 11-346863

PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: JP 11-34684

PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: JP 2000-31062

PRIOR FILING DATE: 2000-02-08

PRIOR APPLICATION NUMBER: JP 2000-34091

PRIOR FILING DATE: 2000-02-10

PRIOR APPLICATION NUMBER: JP 2000-34090

PRIOR FILING DATE: 2000-02-10

PRIOR APPLICATION NUMBER: JP 2000-35829

PRIOR FILING DATE: 2000-02-14

PRIOR APPLICATION NUMBER: JP 2000-35899

PRIOR FILING DATE: 2000-02-14

PRIOR APPLICATION NUMBER: JP 2000-71161

PRIOR FILING DATE: 2000-03-14

PRIOR APPLICATION NUMBER: JP 2000-160851

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 160

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 82

LENGTH: 233

TYPE: PRT

ORGANISM: Homo sapiens

US-09-890-688-82

Query Match 9.4%; Score 113; DB 3; Length 233;

Best Local Similarity 27.1%; Pred. No. 0.0075;

Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

Db 31 ASFTTSMEDTQVYKSSPLGPGAGPQAPSWLQPERCAVFOCAQCHAVLADS 90

Db 49 ASWSSMSSEDAVY---ADMERQQL--EEHAAAE-----ERPLVFLCSGCRRLPGDS 95

Db 91 VHLAMPISR-SLGAIVFSSVTNNVLEAPFIVIGISLSTYNLIFCCSGCGIPVGFHY 149

Db 96 --LSWASQEDTNCILLRCSVSVDEKQKSKREKNGCVLETICAGCSINLGVYR 153

Db 150 STAAALALRGHFCSSDKWVCYL--KTKAIVNASEMDIONVPLSEKIAELKIVLTH 207

Db 154 CTPKNDYKRDLCFLSVEALIESVVGSSSEKQIV-SDKELFNL---ESRVEIKSLTQME 209

OY 208 NRKSLMKLISE 219  
Db 210 DVLKALQMKLWE 221

RESULT 8  
US-10-408-765A-969  
Sequence 969, Application US/10408765A  
Publication No. US20040101874A1  
GENERAL INFORMATION:  
APPLICANT: Ghosh, Soumitra S.  
APPLICANT: Fahy, Brian D.  
APPLICANT: Zhang, Bing  
APPLICANT: Gibson, Bradford W.  
APPLICANT: Taylor, Steven W.  
APPLICANT: Glenn, Gary M.  
APPLICANT: Marnock, Dale E.  
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
FILE REFERENCE: 660088.465  
CURRENT APPLICATION NUMBER: US/10/408,765A  
CURRENT FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 3077  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 969  
LENGTH: 233  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-408-765A-969

Query Match 9.4%; Score 113; DB 4; Length 233;  
Best Local Similarity 27.1%; Pred. No. 0.0075;  
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;  
OY 31 ASFTTSMEDTQVYVSSPLQAGAEPPAAGPOLPEWLPBCAVFOCAQCHAVLADS 90  
Db 49 ASMSSSESDSV--ADMERALQ--EEBAALAE-----ERPLVLGCGCRRPLGDS 95  
OY 91 VHLAMDLSR-SLGAVVFSRVNTNNVLEAPFLVIGSLKSGSTYMLFCGSGGIPVGFHLV 149  
Db 96 --LSWVVSQEDPTNCLRCVSCNVSDVDEKQKLSKREKENGCVLETLCCAGCSLNLGYR 153  
OY 150 STHAALALRGHFCISSDKMVCYLL--KTALVNASEMDIQVPLSEKIAELKEKIVLTH 207  
Db 154 CTPKRLDYKRLFCLSVEAIESYVIGSGSEKQIV-SEDEKELFNL--ESRVEIEKSLTQWE 209  
OY 208 NRKSLMKLISE 219  
Db 210 DVLKALQMKLWE 221

RESULT 9  
US-10-437-963-195427  
Sequence 195427, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 195427  
LENGTH: 708

TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_91377C.1.pep  
US-10-437-963-195427

Query Match 8.0%; Score 95.5; DB 4; Length 708;  
Best Local Similarity 24.5%; Pred. No. 2.1;  
Matches 71; Conservative 32; Mismatches 86; Indels 101; Gaps 17;  
OY 6 LRHRRCATPPRGDF-----CGTERPAIDQASTTSMEW 39  
Db 225 LRHRRCATPSADSPPPRQSVITIGEKARAKAARAKSGGTSAA-----SPTVST 279  
OY 40 DTQVYVSSPLGAGAEPPAAGPOLP--SW--LQPCAVF-----QCAQCH 84  
Db 280 DVPPVVSQEVTPSG-PISDPAGPSLPEAVLTWEELOVEMGRLLKAGANGREISBAR 338  
OY 85 AVLADSVHLAMDLSRSLG-----AVVFSRVNTNNVLE-----APFLVIGESLKL 128  
Db 339 AETPAALALRELRLVLEAREBDLTKRELVAAGNERQSKLEDMSELGDLSEIRGSLR 398  
OY 129 GSTYNNLL-FCGSGG--IPVG--FHLVSTHAALALRG--HFCUSSDKM----- 169  
Db 399 -VTYTGHLQAKEGVKSTIPVNLDEPSTLSLAELATAMGEIIPSKHTSRIAERTSNGIV 457  
OY 170 --VCYLLKTALVNAS--EMDIQNV-----PLSEKIAELKEKIV 204  
Db 458 TGVCHVL--ACVRLSRPELDRLREILDGAASDTRKEWMEVSDLSGSLV 504

RESULT 10  
US-10-437-963-154548  
Sequence 154548, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 154548  
LENGTH: 361  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_54398C.1.pep  
US-10-437-963-154548

Query Match 7.7%; Score 92; DB 4; Length 361;  
Best Local Similarity 25.3%; Pred. No. 1.9;  
Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;  
OY 2 AAGPLRHRRCATPPRGDFCGTERAIDQASTTSMEDTQVYVSSPLGAGAEPPA 61  
Db 49 APTLPRRAVRAKAAODGSGTSSASPAAVAST-----DVVVVSGSRATPSG-PASDPV 102  
OY 62 AGPOLP-----SW--LQPCAVFQCAQCHAVLADSVHLAMDLSRSLGAVVFSRVNTNNV 115  
Db 103 AGRGSAAVLSWEELOVEMGRLLBAGA--RVIGREIAPARLEHRM-----SELGN--- 152  
OY 116 EAPFLVIGESLKGSTYNNLL--FCGSGGI-----PVGFHLVSTHAALAL--RGH 161  
Db 153 -----LSEIRGSLR-VTYTGHLQLAGKGIGKSTIPANPDEFSLTSLAELAAEMERIPSKH 207

QY 162 FCLSDKN-----VCYLTKTAIVNASEMDIONV-----PLSEKIAELKEKI 203  
Db 208 AARIGEMSNRIYIGACHILACVRLAH-PELDREILQGEASDARKDQWESVGLGRSV 266  
QY 204 V 204  
Db 267 L 267

## RESULT 11

US-10-369-493-18389  
Sequence 18389, Application US/10369493  
Publication No. US2003023675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 18389  
LENGTH: 695  
TYPE: PRT  
ORGANISM: Lactococcus lactis  
US-10-369-493-18389

Query Match 7.7%; Score 92; DB 4; Length 695;  
Best Local Similarity 25.3%; Pred. No. 4.7; Indels 50; Gaps 7;  
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;

QY 101 LGAVFSRYTNVNVLEAPLVGIEGSLKSTYNLLFCGSCGIPVGFHLVSTHAALALNG 160  
Db 101 LGTIIF-----FYSGTFFPGAKGELSKRPMWMLIMGITVAV-AVSVAATIMSNG 153  
QY 161 HF-----CLSDKNVCYLTKTAIVNASE--MDI-----QNPVSE-K 195  
Db 154 HMGWFMFELATLIVIMLIGHLEMKAINAGDALDGLASLVKKAHLKSGDVELSEIK 213  
QY 196 IAELEKEKIVLTNNRKLKSLMKIISVTPDOS 225  
Db 214 VGDL--LVKENEKIPADGLILSEALVDES 241

## RESULT 12

US-09-780-525-2  
Sequence 2, Application US/09780525  
Patent No. US2002004223A1  
GENERAL INFORMATION:  
APPLICANT: Bin-Bing Zhou  
APPLICANT: Yuan Zhu  
APPLICANT: Priya Chaturvedi  
APPLICANT: Mark R. Hurtle  
APPLICANT: Xiaolong Li  
TITLE OF INVENTION: PHAR1, A NEW RING FINGER PROTEIN  
FILE REFERENCE: GP-70668-C1  
CURRENT APPLICATION NUMBER: US/09/780,525  
CURRENT FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 09/456,876  
PRIOR FILING DATE: 1999-12-08  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 664  
TYPE: PRT  
ORGANISM: HOMO SAPIENS

## US-09-780-525-2

Query Match 7.4%; Score 89; DB 3; Length 664;  
Best Local Similarity 23.0%; Pred. No. 8.9; Indels 74; Gaps 11;  
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GATERAIDQ-----SFTSMEMDPQVVGKSSPL----- 50  
Db 434 YRQMAQDPH---CPAPGEFGAPQALDADSTSVLTTAVQDYVCPQSGHALCTCGQ 490  
QY 51 ---GPAGLGAEPAPAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLANDLSR----- 99  
Db 491 PMPBRARERQDPRAPO-----QCAVC---LQPFCHLYMGCTRTGCGCIA 534  
QY 100 ---SLCAVVPKRYTNVNVLEAPLVGIEGSLKSTY-NLLFCGSCGIPVGFHLVSTHA 153  
Db 535 PFCEINLDGKCLDGLVNNNSYESDILKNYLAT-RGLTWKNNL-----TBS 578  
QY 154 ALAALRGHFCLS-----SDKNVCYLTKTAIVNASEMDIONVPLSE 194  
Db 579 LVALQGVFLSDYKRVTDVTLVLCYCGLRSRRELITYROYRONIPASE 624

## RESULT 13

US-10-437-963-186569  
Sequence 186569, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 186569  
LENGTH: 435  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT453d\_83357C.1.pep  
US-10-437-963-186569

Query Match 7.4%; Score 88.5; DB 4; Length 435;  
Best Local Similarity 22.9%; Pred. No. 5.6; Indels 59; Gaps 12;  
Matches 50; Conservative 27; Mismatches 82; Indels 59; Gaps 12;

QY 2 AAOPL--RH--RSRCATPPRGDFCGTERAIDQASFTSMEMDPQVVGKSSPLGAGLA 57  
Db 41 APEPLSCRGRHLRCAVD---GAGKRTERRPAPQ-----RBSGSGSLGAL 88  
QY 58 EEPAPGOLPSWLOPERCAVFOCAQCH---VLADSVHLANDLSRSLGAVVFSRYTNVVL 115  
Db 89 EDPSPQPGVPLILP-----LCRCYAKGISEVYVTRTDLVNHL-----LNSNAIS 135  
QY 116 EAPFLVIGSLKSTYNLLFCGSCGIPVGFHLVSTHAALALRGHFCLSDDKMYGLAK 175  
Db 136 EGPFF-----SWRKARFLGSAS--AFVVKQTEWP-----CATDSSKCY-LQ 173  
QY 176 TKAIYNASEMDION-----VPLSEKIAELKEKIVLTH 207  
Db 174 NSFGGITEDEOSLYNPLPSKELLPDDKMSIFDH 211

## RESULT 14

US-10-437-963-155606







APPLICANT: Yang, Yan Ping  
APPLICANT: Klein, Michel H.  
TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE  
FILE REFERENCE: 1038-921MIS:jb  
CURRENT APPLICATION NUMBER: US/11/013,759  
PRIORITY FILING DATE: 2004-12-16  
PRIORITY FILING DATE: 1999-07-27  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 3  
LENGTH: 1992  
TYPE: PRT  
ORGANISM: Moraxella catarrhalis  
US-11-013-759-3

Query Match 5.6%; Score 67.5; DB 7; Length 1992;

Best Local Similarity 22.0%; Pred. No. 18;  
Matches 40; Conservative 20; Mismatches 47; Indels 75; Gaps 7;

QY 46 GSSPLGPAAGLAEPPAAGPOLPSWLPQPCAVFQCAQCHAVLADSVHLAMDLSRLGAVV 105  
DB 964 GKSTLNDGSLIKNPTGSEQI-----QVGADGVK----- 992  
QY 106 FSRVTNNVLEAPFLVIGESLKGSTYNLLFCGSGGIVGFHLVSTHAALALRGHFLS 165  
DB 993 FAKVNNNGVGA---GIDGTRIRTRDEIGFTGNG-----S 1025  
QY 166 SDKWVCYLTKTKAIVNASEMDIONVPLSE-----KIAELKEKIVLTNRRLKSL 213  
DB 1026 LDKSKPHL--SKDGINAGKKITNIOGSEIAQNSHDAVGTGKIYDLKTEL---ENKISST 1080  
QY 214 MK 215  
DB 1081 AK 1082

RESULT 3  
US-11-013-759-13

Sequence 13, Application US/11013759  
Publication No. US20050249747A1  
GENERAL INFORMATION:  
APPLICANT: Loomore, Sheena M.  
APPLICANT: Sasaki, Ken  
APPLICANT: Yang, Yan Ping  
APPLICANT: Klein, Michel H.  
TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE  
FILE REFERENCE: 1038-921MIS:jb  
CURRENT APPLICATION NUMBER: US/11/013,759  
PRIORITY FILING DATE: 2004-12-16  
PRIORITY FILING DATE: 1999-07-27  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 13  
LENGTH: 1992  
TYPE: PRT  
ORGANISM: Moraxella catarrhalis  
US-11-013-759-13

Query Match 5.6%; Score 67.5; DB 7; Length 1992;

Best Local Similarity 22.0%; Pred. No. 18;  
Matches 40; Conservative 20; Mismatches 47; Indels 75; Gaps 7;

QY 46 GSSPLGPAAGLAEPPAAGPOLPSWLPQPCAVFQCAQCHAVLADSVHLAMDLSRLGAVV 105  
DB 964 GKSTLNDGSLIKNPTGSEQI-----QVGADGVK----- 992  
QY 106 FSRVTNNVLEAPFLVIGESLKGSTYNLLFCGSGGIVGFHLVSTHAALALRGHFLS 165  
DB 993 FAKVNNNGVGA---GIDGTRIRTRDEIGFTGNG-----S 1025

QY 166 SDKWVCYLTKTKAIVNASEMDIONVPLSE-----KIAELKEKIVLTNRRLKSL 213  
DB 1026 LDKSKPHL--SKDGINAGKKITNIOGSEIAQNSHDAVGTGKIYDLKTEL---ENKISST 1080  
QY 214 MK 215  
DB 1081 AK 1082

RESULT 4  
US-11-013-759-4

Sequence 4, Application US/11013759  
Publication No. US20050249747A1  
GENERAL INFORMATION:  
APPLICANT: Loomore, Sheena M.  
APPLICANT: Sasaki, Ken  
APPLICANT: Yang, Yan Ping  
APPLICANT: Klein, Michel H.  
TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE  
FILE REFERENCE: 1038-921MIS:jb  
CURRENT APPLICATION NUMBER: US/11/013,759  
PRIORITY FILING DATE: 2004-12-16  
PRIORITY FILING DATE: 1999-07-27  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 4  
LENGTH: 2047  
TYPE: PRT  
ORGANISM: Moraxella catarrhalis  
US-11-013-759-4

Query Match 5.6%; Score 67.5; DB 7; Length 2047;  
Best Local Similarity 22.0%; Pred. No. 18;  
Matches 40; Conservative 20; Mismatches 47; Indels 75; Gaps 7;

QY 46 GSSPLGPAAGLAEPPAAGPOLPSWLPQPCAVFQCAQCHAVLADSVHLAMDLSRLGAVV 105  
DB 1019 GKSTLNDGSLIKNPTGSEQI-----QVGADGVK----- 1047  
QY 106 FSRVTNNVLEAPFLVIGESLKGSTYNLLFCGSGGIVGFHLVSTHAALALRGHFLS 165  
DB 1048 FAKVNNNGVGA---GIDGTRIRTRDEIGFTGNG-----S 1080  
QY 166 SDKWVCYLTKTKAIVNASEMDIONVPLSE-----KIAELKEKIVLTNRRLKSL 213  
DB 1081 LDKSKPHL--SKDGINAGKKITNIOGSEIAQNSHDAVGTGKIYDLKTEL---ENKISST 1135  
QY 214 MK 215  
DB 1136 AK 1137

RESULT 5  
US-11-013-759-7

Sequence 7, Application US/11013759  
Publication No. US20050249747A1  
GENERAL INFORMATION:  
APPLICANT: Loomore, Sheena M.  
APPLICANT: Sasaki, Ken  
APPLICANT: Yang, Yan Ping  
APPLICANT: Klein, Michel H.  
TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE  
FILE REFERENCE: 1038-921MIS:jb  
CURRENT APPLICATION NUMBER: US/11/013,759  
PRIORITY FILING DATE: 2004-12-16  
PRIORITY FILING DATE: 1999-07-27  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7  
LENGTH: 2047  
TYPE: PRT  
ORGANISM: Moraxella catarrhalis  
US-11-013-759-7

Query Match 5.6%; Score 67.5; DB 7; Length 2047;  
Best Local Similarity 22.0%; Pred. No. 18;  
Matches 40; Conservative 20; Mismatches 47; Indels 75; Gaps 7;

QY 46 GSSPLGPAIGAEPAAGPOLPSWLPQPERCAVFCQAQCHAVLADSVHLAMDLSRSLGAV 105  
DB 1019 GKSTINDGSLSTKXNPTGSEQI-----QVADSVK----- 1047  
QY 106 FSRVTNVVLEAPFLVIGSLKSTYVLLFCGSCGIPVGFHLYSTHAALALRGHPCIS 165  
DB 1048 FAKVNNNGVVG-----GIDGTRITRDELIGFTGTG-----S 1080  
QY 166 SDRKVCYLTKTKAIVNASMDIQVPLSE-----KIALKEKIVLTHRLKSL 213  
DB 1081 LDKSKPHL--SKDGINAGGKITNTIQSGEIAQNSHDVATGKTYDLKTEL---ENKISST 1135  
QY 214 MK 215  
DB 1136 MK 1137

RESULT 6  
US-10-689-742-172  
Sequence 172, Application US/10689742  
Publication No. US20050250180A1

GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John M  
APPLICANT: Lavalie, Edward R  
APPLICANT: Racie, Lisa A  
APPLICANT: Evans, Cheryl  
APPLICANT: Weiberg, David  
APPLICANT: Treacy, Maurice  
APPLICANT: Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
FILE REFERENCE: 00766.000091.10  
CURRENT APPLICATION NUMBER: US/10/689,742  
CURRENT FILING DATE: 2003-10-22  
PRIOR APPLICATION NUMBER: 09/746,783  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 231  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 172  
LENGTH: 381  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-689-742-172

Query Match 5.6%; Score 67; DB 1; Length 381;  
Best Local Similarity 21.7%; Pred. No. 1.8;  
Matches 23; Conservative 19; Mismatches 42; Indels 22; Gaps 4;

QY 78 PQCAQCH-----ATLADSVHL-----AMDLSRSLGAVFVSVTNNVLEAPFLVIGSLG 125  
DB 202 FKPQNNKREPFQENLRMGITHIPDRDAWELSPGAFSDLYQYQHCADAPICPEQGRDS 261  
QY 126 SLKSTYVLLFCGSCGIPVGFHLYSTHAALALRGHFCISDPMVC 171  
DB 262 FEDEGWCILLCATCG-----SHGTHRDCSSLR-----FNSKMKWC 297

RESULT 7  
US-10-336-263A-8  
Sequence 8, Application US/10336263A  
Publication No. US20050251882A1  
GENERAL INFORMATION:  
APPLICANT: D'Ordine, Robert L

APPLICANT: Dotson, Stanton B  
APPLICANT: Duff, Stephen M  
APPLICANT: Sisson, Pamela J  
TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE NUCLEIC ACID MOLECULES FROM CYANOBACTE  
FILE REFERENCE: 38-21 (52176) A  
CURRENT APPLICATION NUMBER: US/10/336,263A  
CURRENT FILING DATE: 2003-01-03  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 480  
TYPE: PRT  
ORGANISM: Nostoc punctiforme  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(480)  
OTHER INFORMATION:

US-10-336-263A-8

Query Match 5.5%; Score 66; DB 1; Length 480;  
Best Local Similarity 28.9%; Pred. No. 3.3;  
Matches 22; Conservative 14; Mismatches 30; Indels 10; Gaps 3;

QY 149 YSTHAALALRGHF-CLSSDKMVCYLTKTKAIVNASMDIQVPLSEKIAELKIVLTH 207  
DB 407 PSTQGVARVHELYQSTLTDFTIQETIKTKLPLERQIQNL-----LKS-K-VLKS 457  
QY 208 NRLKSLMKILSEVTPD 223  
DB 458 NEIKALEKLIQSFMSND 473

RESULT 8  
US-10-467-962B-101  
Sequence 101, Application US/10467962B  
Publication No. US20050246784A1  
GENERAL INFORMATION:  
APPLICANT: Plesch, Gunnar  
APPLICANT: Blau, Astrid  
APPLICANT: Daeschner, Klaus  
APPLICANT: Klein, Mathieu  
TITLE OF INVENTION: Identification of Herbicidally Active Substances  
FILE REFERENCE: 2000 857  
CURRENT APPLICATION NUMBER: US/10/467,962B  
CURRENT FILING DATE: 2003-08-14  
PRIOR APPLICATION NUMBER: PCT/EP02/01466  
PRIOR FILING DATE: 2002-02-13  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: PatentIn Vers. 2.0  
SEQ ID NO 101  
LENGTH: 680  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-10-467-962B-101

Query Match 5.4%; Score 65; DB 1; Length 680;  
Best Local Similarity 24.6%; Pred. No. 7;  
Matches 43; Conservative 25; Mismatches 81; Indels 26; Gaps 8;

QY 73 EBCAVPQCAQCHAVLADSV--HLAW--DLRSLSGAVFVSVTNNVLEAPFLVIGSLG 129  
DB 477 EERGCLPCATYTNIMINISYGDERMDMKRLGKMGSGIIPVVTHTT-LVDVYG--KS 533  
QY 130 STYN-----LFCGSCGIPVGFHLYSTHAALALRG-----HFCSSDKMVCYLTKT 176  
DB 534 GRFNDAIECLEEMKSVGLKPSSTMYNALINAYAGSLSEQAVNAPRVMTSDGLKPSLAL 593  
QY 177 KAIIVNASMDIQVPLSEKIAELKEK---IYLTNRLKSLMKILSEVTPDQSKP 227  
DB 594 NSLINAFGEDRRDAEFAVLYQYMKENGVPDVVYTY--TLMKALIRVDFKQVP 645

```

      1  APPLICANT: Plescu, Gunnar
      2  APPLICANT: Blau, Astrid
      3  APPLICANT: Daeschner, Klaus
      4  APPLICANT: Klein, Mathieu
      5  TITLE OF INVENTION: Identification of Herbicidally Active Substances
      6  FILE REFERENCE: 2000 857
      7  CURRENT APPLICATION NUMBER: US/10/467, 962B
      8  CURRENT FILING DATE: 2003-08-14
      9  PRIOR APPLICATION NUMBER: PCT/EP02/01466
     10  PRIOR FILING DATE: 2002-02-13
     11  NUMBER OF SEQ ID NOS: 109
     12  SOFTWARE: PatentIn Ver. 2.0
     13  SEQ ID NO 95
     14  LENGTH: 592
     15  TYPE: PR1
     16  ORGANISM: Arabidopsis thaliana
     17  US-10-467-962B-95

Query Match
Best Local Similarity 23.0%; Score 62; DB 1; Length 592;
Matches 32; Conservative 22; Mismatches 37; Indels 48; Gaps 7;

OY 103 AVFESRYTNVNIAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAA----- 154
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 325 AVFENRSRRARLQS-----VKCTIAGKNAVYVFCSTGDMGNVWSKGVQNVLEYLTD 379
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 155 -----LAALRGHFCULSSDK-----MVCYLKTRAIYVASEMDIQNVPLSE 194
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 380 DFPDMVDYIGISGNFC--SDKKPAVVMIEGRKSVVC-----EAVING--EIVNKVLKT 429
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 195 KIAELKEKIVLTHNRKSL 213
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 430 SVALVEL-----NMLKNL 443
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 11
US-10-997-201A-30
Sequence 30. Application US/10997201A
Publication No. US20050249739A1
GENERAL INFORMATION:
APPLICANT: Marasco, Wayne
APPLICANT: Su, Jianhua
TITLE OF INVENTION: Antibodies Against SARS-COV and Methods of Use Thereof
FILE REFERENCE: 20363-026
CURRENT APPLICATION NUMBER: US/10/997, 201A
CURRENT FILING DATE: 2004-11-24
PRIOR APPLICATION NUMBER: 60/524, 840
PRIOR FILING DATE: 2003-11-25
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 669
TYPE: PR1
ORGANISM: Homo sapiens
US-10-997-201A-30

Query Match
Best Local Similarity 22.1%; Score 60.5; DB 1; Length 669;
Matches 30; Conservative 22; Mismatches 57; Indels 27; Gaps 6;

OY 29 DQASFTTSM-EMDTQVVKGSSEPLCPAGLGAEEPPAGFQLPWSWLOPERCAVQCACQ----- 83
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 550 DVSDFTSQVSDPKTSEIJDISEPCAFGVSVITPGTN-----ASSEVAVLQDVNCTDVS 603
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 84 HAVLADVTHLAWDLSRLGAVVFSRVNNVNLBAFLVIGIEGSLKSTYNLLFCGSCGIP 143
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 604 TAIADQDTTPAMR-----IYSTGNVVFOTQAGCLIGAEH--VDTSY-----ECDIP 647
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 144 VGPHLYSTHAAALALR 159
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 648 IGAGICASVHTVSLLR 663
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

## RESULT 12

US-11-074-176-80

Sequence 80, Application US/11074176  
Publication No. US20050250135A1

## GENERAL INFORMATION:

APPLICANT: Kjaenhammer, Todd R.  
APPLICANT: Russell, William M.  
APPLICANT: Altermann, Eric  
APPLICANT: McAuliffe, Olivia  
APPLICANT: Perill, Andrea Azcarate  
TITLE OF INVENTION: Nucleic Acid Sequences Encoding  
Stress-Related Proteins and Uses Therefore  
FILE REFERENCE: 5051-694  
CURRENT APPLICATION NUMBER: US/11/074,176  
CURRENT FILING DATE: 2005-03-07  
PRIOR APPLICATION NUMBER: 60/551,161  
PRIOR FILING DATE: 2004-03-08  
NUMBER OF SEQ ID NOS: 381  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 80  
LENGTH: 614  
TYPE: PR1  
ORGANISM: Lactobacillus acidophilus  
US-11-074-176-80

## Query Match

Best Local Similarity 5.0%; Score 60; DB 7; Length 614;  
Best Local Similarity 19.9%; Pred. No. 21;

Matches 46; Conservative 33; Mismatches 100; Indels 52; Gaps 9;

QY 25 EBAIDASFTTSMEMTQVVKSSPLGPAGLGAEEPAAGPOLPSWQPERCAVFOCAQCH 84

DB 292 DNALKDAGLTVN-DIDKVLNGSTRIPAVQKVKEMAKKEPDHSINPEAVAGAA--- 347

QY 85 ATLASDVHLAMDLSRSLGAVVFSRYTNVYLB-APLVGIEGSLKSTYNLLFCGSCGIP 143

DB 348 -----IQGVISDVAQDYLVDVTPISLGIH-TMGVFTKLIDRNTTIP 390

QY 144 VG-FHLYSTHA-----ALALRGHFCLSDDK-----MVCYLK 175

DB 391 TSKSGFSTADNQPADVHVQLGGERPMADDKTLGRFELTDIPAPRGVQIQVTFDID 450

QY 176 TTAIVNASEMDIONVPLSEKIAELKEKIVLTHRLSLKILSEVTPDQSK 226

DB 451 KNGIVNVSAKM-GTGEQKIT-IKSSGSLSDIEIKRMQDAEHAEDDK 499

## RESULT 13

US-11-074-176-222

Sequence 222, Application US/11074176  
Publication No. US20050250135A1

## GENERAL INFORMATION:

APPLICANT: Kjaenhammer, Todd R.  
APPLICANT: Russell, William M.  
APPLICANT: Altermann, Eric  
APPLICANT: McAuliffe, Olivia  
APPLICANT: Perill, Andrea Azcarate  
TITLE OF INVENTION: Nucleic Acid Sequences Encoding  
Stress-Related Proteins and Uses Therefore  
FILE REFERENCE: 5051-694  
CURRENT APPLICATION NUMBER: US/11/074,176  
CURRENT FILING DATE: 2005-03-07  
PRIOR APPLICATION NUMBER: 60/551,161  
PRIOR FILING DATE: 2004-03-08  
NUMBER OF SEQ ID NOS: 381  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 222  
LENGTH: 639  
TYPE: PR1  
ORGANISM: Lactobacillus acidophilus  
US-11-074-176-222

Query Match 4.9%; Score 58.5; DB 7; Length 639;  
Best Local Similarity 22.5%; Pred. No. 33;

Matches 56; Conservative 31; Mismatches 71; Indels 91; Gaps 17;

QY 25 EBAI--DQASFTTSMEMD-----TQVVKSSPLGPAGLGAEEPAAGPOLPSW 69

DB 160 ERGIYTSIGAFSGAIGMGLAIIVPLVTGYVAVTGNK-----EEGA---PEW 205

QY 70 LQPER-----CAVPOC---AQCNAVLDSDV-----HLAMDLSRSL 101

DB 206 LAFPAVISAALITICAIIVCFGRKEKNLIRNSAKQKTLRQVFSALFNHDQILMP---SL 262

QY 102 GAVVES---RYTNVY-----VLBAP---FLVIGIEGSLKSTYNLLFCGSCGIPVGFHL 148

DB 263 ALVFSLSANTITNGVLFYLYKFKVIGKPGEFVWGLIATIG-----FCVSPMFP1-LNK 315

QY 149 YSTHAALALRGHFCISDQKVC-YLTKKAIWNASMDIONVPLSEKIAELKEKIVLT- 206

DB 316 YIPRKWL-FIACQTC-----MICAVYLFIICRDNVVLDGLILFINFAQLVTLVLTLD 369

QY 207 ---HNRLKS 212

DB 370 AIEYGLKS 378

## RESULT 14

US-10-512-184-27

Sequence 27, Application US/10512184  
Publication No. US2005024901A1

## GENERAL INFORMATION:

APPLICANT: Fraunhofer Gesellschaft zur F"orderung der angewandten Forschung e.V.  
TITLE OF INVENTION: antibody fragments, recombinant  
TITLE OF INVENTION: antibody fragments and fusions mediated plant disease  
FILE REFERENCE: 3581.01US01  
CURRENT APPLICATION NUMBER: US/10/512,184  
CURRENT FILING DATE: 2004-10-22  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 27  
LENGTH: 250  
TYPE: PR1  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: scFv SGB3 with  
OTHER INFORMATION: specificity against Fusarium spp.; originates from  
OTHER INFORMATION: Gallus gallus.  
US-10-512-184-27

Query Match 4.8%; Score 58; DB 1; Length 250;

Best Local Similarity 24.5%; Pred. No. 9.5; Mismatches 27; Conservative 15; Mismatches 40; Indels 28; Gaps 5;

QY 22 GGTGAIDQASFTTSMEMTQVVKSSPL-----GPAGT-----GAEEPAAGPOLPS 68

DB 156 GGTIVK-ITSGSGFAHYSWQKSPGSAVPTLISFNNQRSDIPSRFSGSKSGTGTLTIT 214

QY 69 WLOPERCAVFOCAQCAVAVLDVSLAMDLSRSLGAVVFSRYTNVYLBAP 118

DB 215 GVAEADNAVYCG-----GWD--RSITAGLFAGTTLTVLQGP 250

## RESULT 15

US-10-131-826A-302

Sequence 302, Application US/10131826A  
Publication No. US20050245730A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C128
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 302
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-302

```

```

Query Match      4.8%; Score 58; DB 1; Length 421;
Best Local Similarity 32.6%; Pred. No. 20;
Matches 28; Conservative 9; Mismatches 27; Indels 22; Gaps 6;

QY      141 GIPVGFHLYSTHAALALRGHCLSS--DKWVCYLKTKAIVNASEMDIQNVPLSEKIAE 198
      52 GVP--FHGYT---LGCVSECPCPTNFPSSWYCDNRKLTIPNI-PMHIQQLYIQFNEIE 104
      199 -----LKEKIVLTNRRLKS 212
      105 AVTANSFINATHLKE-INLSHNRKIKS 129
DB

```

Search completed: November 18, 2005, 20:49:50  
 Job time : 8 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: November 18, 2005, 20:27:55 ; Search time 40 Seconds  
(without alignments)  
550.841 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198  
Sequence: 1 MAAPLRHRRCATPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62  
Gapox 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	7.7	695	2 C86731	copper-potassium t
2	90.5	7.6	597	2 D71293	probable phosphori
3	89	7.4	306	2 T46399	hypothetical prote
4	88	7.3	250	2 A43623	kanamycin kinase (
5	86.5	7.2	669	2 A97229	NAD-dependent DNA
6	85	7.1	2194	1 JQ1977	glutamate synthase
7	84.5	7.1	361	2 AE2074	ferrichrome bindin
8	83	6.9	663	2 D97047	DNA ligase (NAD de
9	82.5	6.9	143	2 B69099	formate hydrogenly
10	82.5	6.9	493	2 A85433	sugar transporter
11	82	6.8	967	2 F87678	DNA polymerase I (
12	81.5	6.8	329	2 D90404	transport protein,
13	81.5	6.8	457	2 H85677	glucuronide permea
14	81.5	6.8	457	2 C90919	glucuronide permea
15	81.5	6.8	457	2 B64918	glucuronide permea
16	81.5	6.8	640	2 T41977	hypothetical prote
17	81.5	6.8	662	2 T44221	probable capsid pr
18	81	6.8	1418	2 S40764	hypothetical prote
19	80.5	6.7	766	2 B85440	receptor kinase-11
20	80.5	6.7	1711	1 A55148	protease-tyrosine-P
21	80	6.7	469	2 S55167	IME2-dependent sig
22	79	6.6	386	2 H90789	probable antimetab
23	79	6.6	386	2 C85650	probable antimetab
24	79	6.6	437	2 C86823	GTP-binding protei
25	78.5	6.6	188	2 F8316	late competence op
26	78.5	6.6	578	2 F86484	probable hydroxyme
27	78.5	6.6	697	1 QOBRK3	U104 protein - hu
28	78	6.5	351	2 H82098	UDP-3-O-3-hydroxym
29	77.5	6.5	662	2 T44036	hypothetical prote

30	77.5	6.5	1013	2 T31211	trnc protein homol
31	77.5	6.5	1643	2 T14274	versican precursor
32	77	6.4	381	2 AD1113	hypothetical prote
33	77	6.4	721	2 F87611	tonB-dependent rec
34	77	6.4	738	1 TFRHM	melanotransferrin
35	76.5	6.4	307	2 AG2017	glycerol-3-phospha
36	76.5	6.4	387	2 D69392	probable acyl-CoA
37	76.5	6.4	380	2 S57526	cellulase - Fibrob
38	76.5	6.4	764	2 H98143	cbdbc protein (u60
39	76.5	6.4	764	2 AD3144	formate dehydrogen
40	76.5	6.4	1554	2 T06370	probable DNA (cyto
41	76	6.3	367	2 AH0936	glycerol dehydroge
42	75.5	6.3	310	2 A55053	endothelial monocy
43	75.5	6.3	415	2 T46716	hypothetical prote
44	75.5	6.3	700	2 D70951	probable UvrD - My
45	75.5	6.3	1534	2 S59604	DNA (cytosine-5-)-

#### ALIGNMENTS

RESULT 1  
C86731  
copper-potassium transporting ATPase B copB [imported] - Lactococcus lactis subsp. lactis  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: C86731  
R:Boletini, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.  
A:Reference number: AB6625; MUID:21235186; PMID:11337471  
A:Accession: C86731  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-695 <STO>  
A:Cross-references: UNIPROT:O9CH87; UNIPARC:UPI00000C6919; GB:AE005176; PDB:G12723778; E  
A:Experimental source: strain IL1403  
A:Genetics:  
A:Gene: copB  
A:Superfamily: Enterococcus copper-transporting ATPase copB; ATPase nucleotide-binding

Query Match 7.7%; Score 92; DB 2; Length 695;  
Best Local Similarity 25.3%; Pred. No. 2;  
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;

QY 101 LGAVVFSRVTVNNVLEAPLVGIEGSLKSTYLLFCGSCGIPVGFHLYSTHAALALRG 160  
DB 101 LGTTIF-----FYSCTPFPSGAKGELKSRKPMMLITMGITVAV-AYSVATIMSLNG 153  
QY 161 HF-----CLSDRMVCYLTKTKATYNASE--MDI-----QNVPISE-K 195  
DB 154 HNGMFWFELATVILMGLGHLEMKATMGADDAKDIALSLVPRKKAHLKSGDVELSEIK 213  
QY 196 IAELEKIVLTNRRLKSLMKILSEVTPDQ 225  
DB 214 VGDLL-LVKEKIPADGLILSEALVDES 241

RESULT 2  
D71293  
probable phosphoribosylglycinamide formyltransferase - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: D71293  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwlr  
son, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McD  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98312770; PMID:9665876  
A:Accession: D71293  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A;Residues: 1-597 <COL>  
A;Cross-references: UNIPROT:O83693, UNIPARC:UPI00000D3290, GB:AB001243, GB:AB000520, NID  
A;Experimental source: strain Nichols  
C;Genetics:  
A;Gene: TP0695

Query Match 7.6%; Score 90.5; DB 2; Length 597;  
Best Local Similarity 28.0%; Pred. No. 2.2;  
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;

QY 12 CATPRGDFCGGTERAID---QASFT-----TSNEMDTQVVKSSPLGPA 53  
DB 90 CALP--GHRLEATKNAIDDKTMRACFTRARLRCPRTFLEPDSFAMDT-----PGHA 140  
QY 54 GLGAEPPAAGPQLPSWLOP-ERCAVFOC--ACCAVLAADSVHLMWLSRSLGAVVFSRVY 110  
DB 141 RLCSHLHSAGLSFPLVVKFPTDMGARCGTLAOCXDTLIMACAVARQPSRS----- 190  
QY 111 NNVTLEAPLVIGISLKGSTYNLLFCGSCGIPVGFHLYSTHALAALRGHFLCLSDPMV 170  
DB 191 GRVLE-ETIVGREFSLEG---LIFDGT-----LYVT--ALA-----DRHI 225

QY 171 CY 172  
DB 226 CF 227

RESULT 3  
146399  
hypochemical protein DKFZP434N2420.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: 146399  
R;Ottensmeyer, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23031  
A;Accession: 146399  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-306 <AAA>  
A;Cross-references: UNIPROT:Q96EP1, UNIPARC:UPI000006D779, EMBL:AL137561  
A;Experimental source: adult testis; clone DKFZP434N2420  
C;Genetics:  
A;Note: DKFZP434N2420.1

Query Match 7.4%; Score 89; DB 2; Length 306;  
Best Local Similarity 23.0%; Pred. No. 1.3;  
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPRGDFC-----GTERAIDQA-----SFTSMEMDTQVVKSSPL----- 50  
DB 76 YRRQAAQPPH--CPAPGEPAPQALDAPSTSVSLTTAVQDYVCPLOGSHALCTCCFQ 132  
QY 51 ---GPAGLGAEPAPQPLPSWLOP-ERCAVFOCAQCAVLAADSVHLMWLSR----- 99  
DB 133 PMPRRAREREDPRAPQ-----QCANC---LQPFCLLYGCMRTGCGCLA 176  
QY 100 ----SLGAVVFSRVYNNVLEAPLVIGISLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153  
DB 177 PCCEINLDKCDLGVANNNSYESDILKNYLAT-RGLTWKNNL-----TES 220  
QY 154 ALAALRGHFLCS-----SDKWVCYLLKTKAIVNASSEMIONVPLSE 194  
DB 221 LVALLRGVFLSDYRVVTGDTVLCGCCGLRSFRELTYQVRONIPASE 266

RESULT 4  
A43623  
kanamycin kinase (EC 2.7.1.95) - Campylobacter jejuni  
C;Species: Campylobacter jejuni  
C;Date: 11-Feb-1993 #sequence\_revision 11-Feb-1993 #text\_change 31-Dec-2004  
C;Accession: A43623  
R;Tenover, F.C.; Gilbert, T.; O'Hara, P.

Plaamid 22, 52-58, 1989  
A;Title: Nucleotide sequence of a novel kanamycin resistance gene, aphA-7, from Campylobacter  
A;Reference number: A43623; MUID:89387451; PMID:2550983  
A;Accession: A43623  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-250 <TEN>  
A;Cross-references: UNIPROT:P14508, UNIPARC:UPI000012DEBC, GB:M29953, GB:U03316, NID:G144  
C;Superfamily: aminoglycoside 3'-phosphotransferase (kanamycin kinase)  
C;Keywords: phosphotransferase

Query Match 7.3%; Score 88; DB 2; Length 250;  
Best Local Similarity 28.7%; Pred. No. 1.3;  
Matches 33; Conservative 18; Mismatches 46; Indels 18; Gaps 5;

QY 105 VFSRVYNNVLEAPLVIGISLKGSTYNLLFCGSCGIPVGFHLYSTHALAALRG-HFC 163  
DB 40 IFSKTYTSVREAEEMMMWLSDKLVDPV-----IEYVGRHSBYLWSELRGKHID 90  
QY 164 LSSDKWVCYLLKTKAIVNA-----SEMDIONVPLESEK-ALKEKIVLTNRKLKL 213  
DB 91 CFIDHPKTYI---ECLVNAHQLOAIDIRNCPSSKIDVRLKELKYLDDRIRADI 142

RESULT 5  
A97229  
NAD-dependent DNA ligase [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 31-Dec-2004  
C;Accession: A97229  
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: A97229  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-669 <KUR>  
A;Cross-references: UNIPROT:Q97FQ5, UNIPARC:UPI00000CA5D7, GB:AB001437, PIDN:AAK80620.1;  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC2673  
C;Superfamily: DNA ligase (NAD), Liga type

Query Match 7.2%; Score 86.5; DB 2; Length 669;  
Best Local Similarity 27.4%; Pred. No. 6.2;  
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8;

QY 96 DLNR---SLGAVVFSRVYNNVLEAPLVIGISLKGSTYNL-----LFCGSCGIPVGF 146  
DB 366 DIGKKYKIGSRVYRNSDVI---PRIMGVTEBTBETNIEAPTCIPYGSISIVKEGV 422  
QY 147 HLVTAAALALRGHFLCLSSDKWVCYLLKTKAIVNASSEMIONVPLESEKIAE-LKEKIV 204  
DB 423 HL-----PC-ENTLISCKPQWKSIVHFASREAMNIEGSEKTAEOLEFEX-- 465  
QY 205 LTHRLKSLMKILSE 219  
DB 466 LNIKISLDLYRTKE 480

RESULT 6  
J01977  
glutamate synthase (NADH2) (EC 1.4.1.14) precursor [validated] - alfalfa  
C;Species: Medicago sativa (alfalfa)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-Oct-2004  
C;Accession: J01977; P00551  
R;Gregerson, R.G.; Miller, S.S.; Twary, S.N.; Gantt, J.S.; Vance, C.P.  
Plant Cell 5, 215-226, 1993  
A;Title: Molecular characterization of NADH-dependent glutamate synthase from alfalfa no.  
A;Reference number: J01977; MUID:93200806; PMID:8453303  
A;Accession: J01977



A:Molecule type: mRNA  
A:Residues: 1-2194 <GRE>  
A:Cross-references: UNIPROT:Q03460; UNIPARC:UPI000012B7FA; GB:L01660; MID:9166411; PIDN:  
A:Accession: P00551  
A:Molecule type: protein  
A:Residues: 102-114 <GR2>  
A:Cross-references: UNIPARC:UPI0000172061  
C:Comment: This enzyme catalyzes the reductive transfer of the amido group of glutamine  
C:Superfamily: glutamate synthase (NADH/NADPH), eukaryotic type  
C:Keywords: 3Fe-4S; chloroplast; iron-sulfur protein; metalloprotein; NAD; oxidoreductase  
F:1-101/Domain: propeptide #status predicted <PRO>  
F:102-2194/Product: glutamate synthase (NADH) #status experimental <MAT>  
F:102/Active site: Cys #status predicted  
F:1246,1252,1257/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 7.1%; Score 85; DB 1; Length 2194;  
Best Local Similarity 24.7%; Pred. No. 39;  
Matches 58; Conservative 22; Mismatches 79; Indels 76; Gaps 12;

4 QPLRRSRCAVPPRGDFCGTERAIDQAS-----FTSMEMDQVVKSSPLGPA 53  
Db 1037 EPLADGSR--NPKRS-----AIKQVASGRFVSSYLTVNADBLQIKMAQAKP----- 1082

QY 54 GLGAEP-----AAGPOLPSWLOPERCAVFCQACQAVLADSVHLAMDLSR 99  
Db 1083 GEGGELPGHKVIGDIATRNSTAGVGLIS--PPPHHDYIS-----IEDLAQLIHDKN 1133

QY 100 SLGAVVPSRVNNVLEAPLVIGISLKGSTYNLLFCG-----SCGIPVG 145  
Db 1134 ANPA--ARISVKLSEAGVGVASGVVGAHEHLISHDGCTGASRWTKSKAGLPME 1190

QY 146 FHLVSTHAALAA--LRGHFCLSSDKMVCYLTKT-----KAIVNASMDIQNVPL 192  
Db 1191 LGLAETHQTVLNDLRGRTTLQTDG---LKTGRDVAITALLGAEEVGFSTAPL 1241

RESULT 7  
AE2074  
ferrichrome binding protein of ABC transporter all2147 [imported] - Nostoc sp. (strain F  
C:Species: Nostoc sp. PCC 7120  
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AE2074  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 6, 205-213, 2001  
A>Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AE2074  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-361 <KUR>  
A:Cross-references: UNIPROT:Q9YV34; UNIPARC:UPI00000CE32E; GB:BA000019; PIDN:BA073846.1;  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all2147  
C:Superfamily: ferrichrome-iron transport protein fecB

Query Match 7.1%; Score 84.5; DB 2; Length 361;  
Best Local Similarity 26.1%; Pred. No. 4.3;  
Matches 36; Conservative 17; Mismatches 36; Indels 49; Gaps 7;

QY 97 LSRSLGAVV-----SVVT--NNVLEAPLVIGISLKGSTYNLLFCGSC-----GIPV 144  
Db 81 ISHALGKVKLPKQKRVVLEENIILDSVTLGK-----PGVWYCCQCEENFRGIP- 133

QY 145 GFHLVSTHAALALRGHFLCSSDKMVCYLKTKRAIVNASMDIQNVPLSEKIALKELKIV 204  
Db 134 -----SD-----LLADVAVVG-----NIGNQGLEKIIISLKPDLI 163

QY 205 LTHNRLSKLMTIISVTP 222  
Db 164 LGLTWLKSXYKILSIAP 181

RESULT 8  
D97047  
DNA ligase (NAD dependent), Liga [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 31-Dec-2004  
C:Accession: D97047  
R:Noelling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daily, M.C.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: D97047  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-663 <KUR>  
A:Cross-references: UNIPROT:Q97US8; UNIPARC:UPI00000CA10F; GB:AE001437; PIDN:AAK79167.1;  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC1195  
C:Superfamily: DNA ligase (NAD), Liga type

Query Match 6.9%; Score 83; DB 2; Length 663;  
Best Local Similarity 28.6%; Pred. No. 13;  
Matches 36; Conservative 18; Mismatches 44; Indels 28; Gaps 7;

QY 102 GAVVPSRVNNVLEAPLVIGISLKGSTYNL-----LFCGSGGIVGFHLVSTHAAL 155  
Db 368 GARVPLRRSNDVI--PRIMGTEETEGTEKEIEAPTCPCGSEIVEGVHL----- 417

QY 156 AALRGHFLCSSDKMVCYLKTKRAIVNASMDIQNVPL-SEKIAE-LKEKIVLTHNRLKSL 213  
Db 418 -----FC--ENTLSCPKQMVKSIVHFASRKAMNIEGSEKTAEDLFEK--LNIKSISDL 467

QY 214 MKTISE 219  
Db 468 YRTTKE 473

RESULT 9  
B69099  
formate hydrogenlyase, iron-sulfur subunit 2 - Methanobacterium thermoautotrophicum (str  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: B69099  
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
J. Bacteriol. 179, 7135-7155, 1997  
Kl, S.; Church, G.M.; Daniele, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: B69099  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-143 <MTN>  
A:Cross-references: UNIPROT:Q27769; UNIPARC:UPI0000066781; GB:AE000929; GB:AE000666; NIT  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTN1736  
C:Superfamily: nrfC protein; ferredoxin 2(4Fe-4S) homology  
F:62-118/Domain: ferredoxin 2(4Fe-4S) homology <FRR2>

Query Match 6.9%; Score 82.5; DB 2; Length 143;  
Best Local Similarity 23.0%; Pred. No. 2.1;  
Matches 35; Conservative 26; Mismatches 66; Indels 25; Gaps 6;

QY 72 PERCAVFCQACQAVLADSVHLAMDLSRSLGAVV-----SRVTNNVLEAPLVIGIEG 125  
Db 8 PELCD--ECWKCEKIRPKKAIRVID-----GVPVFCMGCSPEAPCLNICEPAIVBDG 60

QY 126 SLKGSTYNLLFCGSC--GIPVGFHLVSTHAALALRGHFLCSSDKMVCYLKTKRAIVNAS 183  
Db 126 SLKGSTYNLLFCGSC--GIPVGFHLVSTHAALALRGHFLCSSDKMVCYLKTKRAIVNAS 183



C;Superfamily: melbiose carrier protein

Query Match	6.8%	Score 81.5;	DB 2;	Length 457;
Best Local Similarity	25.2%;	Pred. No. 11;		
Matches 37;	Conservative 30;	Mismatches 53;	Indels 27;	Gaps 8;

QY 95 MDLRSLCAVVFSTRNNV-----VLEA-----PFLVIGIBSLKSGITYNL-FCGSGC 141  
 Db 317 WSLPALVALVALIASGGQVYTMVMALEADYVGEYGYLTGVR--IGLRTSLPSFRKCG 374  
 QY 142 I PVGHLYSTHAALALRGHFC--LSDDKNCYCLKTKAIYNASEMDIONV-----PLGE 194  
 Db 375 QAIG--GSLPFLILGSGYIANOVQTPPEVINGIRTSIALVPCGMLFAVITWFPPLD 431

```

Oy      195 KIAELKEKIVLTNNRLKSLMKILSEVT 221
          | : | | | | : : : |
Db      432 K-KFKIIVEIDNRKKVQQLISDIT 456

```

RESULT 14  
C90919

```
glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05089)
CISpecies: Escherichia coli
C.Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #next_change 09-Jul-2004
```

R. Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.  
Gasegawara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

A1:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genome  
 A1:Reference number: A99629; MUID:2156231; PMID:11258796

A;Accession: C90919

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-457 &lt;HAY&gt;

A; Cross-references: UNIPROT:Q8X673; UNIPARC:UP100000D0BE5; GB:BAU0000//; P1DN:BA635/46.1

A; Experiment  
C; Genetics

A:Gene: EC82323

A/gene: bcb2323  
C/Superfamily: melibiose carrier protein

'Query Match	6.88; Score 81.5; DB 2; Length 457;
--------------	-------------------------------------

```

QY 95 DDLSLSLGLAVVSRVTNNV-----VLEA-----PELVGIGSLGSLGTYNNL-FCGSGC 141
QY 317 WSLPALVALVALAIAISGGQVTMTVMMALEADYVEYGYLLGV-IGLTLTSLSPFRKKC 374
Db 142 IPVGHLVSTHAALALRGHFC--LSSDKVCCYLLETKTAIVNASEMDIONV-----PLSE 194
QY 375 QAIG---GSIPFIFIGLGYIANQVTFPVINGIRTSIALVPCGFMLLFVIMFEYPLD 431
Db

```

QY 195 KIAELKEKIVLTHNRLKSLMKILSEVT 221  
| : | : | : : : : :  
Db 432 K-KFKEIVVEIDNRKKVQOQLISDIT 456

## RESULT 15

glucuronide permease uidB - *Escherichia coli* (strain K-12)

C:Species: Escherichia coli  
C:Date: 13-Sep-1997 #Accession: 17-Sep-1997 #Text change: 09-Jul-2004

C:\Date: 12-Sep-1997  
C:\Accession: B64918

C:\ACCESSION: B64918  
B:Blattner, F.B.: Plunkett III, G.

A.: Rose, D.J.: Mau, B.: Shao, Y.

Science 277, 1453-1456 (2002)

A;Title: The complete genome sequence of *Escherichia coli* K-12.

**A;Reference number:**

A;Accession: B64918

A/Status: nucleic acid sequence not shown; translation not shown

A; Experimental source: strain K-12, substrain MG1655

C;Genetics:

A; Gene: uidb; gusb  
C: Superfamily: mel

C:keyword: carrier protein: transmembrane  
C:superfamily: melibiose carrier protein

F137-53/Domain: transmembrane #status predicted <T

F;81-97/Domain: transmembrane #status predicted <TM02>

F;152-168/Domain: transmembrane #status predicted <TM03>

F;184-200/Domain: transmembrane #status predicted <TM04>

Accession	Protein	Length	Weight	PI	Instability	Transmembrane	status	predicted	<TM05>
F231-247	Domain	16	1.5	4.5	0.0	1	1	1	1

F;263-279/Domains:	transmembrane	#status	predicted	<TM07>
E;310-326/Domains:	transmembrane	#status	predicted	<TM06>

F:408-424/Domain:	transmembrane	#status	predicted	<TM08>
F:310-326/Domain:	transmembrane	#status	predicted	<IM07>

F;908-424/DOLLARD: CLAIMS IN THE PROCESS - 2006

Query Match 6.8%; Score 81.5; DB 2; Le

Query Match	6.8%	Score 81.5;	DB 2;	length 457;
Best Local Similarity	25.2%;	Pred. No. 11;		
Matches 37;	Conservative 30;	Mismatches 53;	Indels 27;	Gaps 8;

```

QY      95  WDLSSLSGLAVVPSRRTNNV-----VLEK-----PELVIGESLSKSGTYNNL-FCGSCG 141
Db      317  WSLPALVALALASIGQGYTMTVMALADIVYEGEYLTVGR--IKGLTSLPSFTRKCG 374
QY      142  IPVGFHYLSTHAALALRGHPC--LSDGMVCYLKTKAIYVNASENDIGNV-----PLSE 194
Db      375  QAIG---GSIPAFILGLSGYIANOVQTPBIVMGIRTSIALVPCGFMLLAFVIITWFPYLT 431
QY      195  KIAELKEKIVLTHNRUKSLMKILISEVT 221
Db      432  K--KEKEIVAEIDNRKKVQOQLSIDT 456

```

Search completed: November 18, 2005, 20:36:43  
Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 20:24:56 ; Search time 232 Seconds  
(without alignments)  
696.405 Million cell updates/sec

Title: US-09-942-052a-728

Perfect score: 1198

Sequence: 1 MAAQPLRRHRCATPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing filter 45 summaries

Database : uniprot\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	1 OIP5_HUMAN	O43482 homo sapien
2	122	10.2	218	2 O9CXR6_MOUSE	O9CXR6 mus musculu
3	118	9.8	462	2 O4PAB5_USITMA	O4PAB5 usitlago ma
4	115.5	9.6	204	2 O9CZJ6_MOUSE	O9CZJ6 mus musculu
5	113	9.4	233	1 CU045_HUMAN	O9Y9P9 homo sapien
6	113	9.4	233	2 O542Z0_HUMAN	O542Z0 homo sapien
7	110.5	9.2	232	1 CU045_PANTR	O68U55 pan troglod
8	109	9.1	207	2 O5B1B1_BRABE	O5B1B1 brachydanio
9	107	8.9	532	2 O54H16_DICDI	O54H16 dictyostei
10	105	8.8	155	1 YCZC_SCHPO	O93802 schizosacch
11	94	7.8	1556	2 O406Z6_LEIMA	O406Z6 leishmania
12	92	7.7	361	2 O86C11_ORISA	O86C11 oryza sativ
13	92	7.7	695	2 O9C8H7_LACIA	O9C8H7 lactococcus
14	91	7.6	628	2 O69KE0_ORISA	O69KE0 oryza sativ
15	90.5	7.6	597	2 O83693_TREPA	O83693 treponema p
16	89.5	7.5	502	2 O50PD2_ENTHI	O50PD2 entamoeba h
17	89.5	7.5	661	2 O50V43_ENTHI	O50V43 entamoeba h
18	89.5	7.5	661	2 O50WZ8_ENTHI	O50WZ8 entamoeba h
19	89	7.4	633	2 O5W052_STRT1	O5W052 streptococc
20	89	7.4	664	1 CHFR_HUMAN	O96E91 homo sapien
21	89	7.4	735	2 O5M4R3_STRT2	O5M4R3 streptococc
22	89	7.4	893	1 POL2_BAMMA	O65339 barley mild
23	88.5	7.4	152	2 O50XW7_ENTHI	O50XW7 entamoeba h
24	88.5	7.4	638	2 O5W052_STRT1	O5W052 streptococc
25	88	7.3	250	1 KKA7_CAMUO	O75766 ciona intes
26	88	7.3	493	2 O68VH2_9POTY	O68VH2 campylobact
27	87.5	7.3	634	2 O6CXH2_KLUTLA	O6CXH2 kluyveromyc
28	87.5	7.3	696	2 O4JRS4_CORUK	O4JRS4 corynebacte
29	87	7.2	664	1 CHFR_MOUSE	O81013 mus musculu
30	86.5	7.2	669	2 O97F05_CLOAB	O97F05 clostridium
31	86.5	7.2	2093	2 O5K2K3_GRHAB	O5K2K3 isthman vir

32	86.5	7.2	5344	2 O529N3_MAGGR	O529N3 magnaportha
33	86	7.2	145	2 O839T4_ENTFA	O839T4 enterococcu
34	86	7.2	327	2 O4NE64_9MTCC	O4NE64 arthrobacte
35	86	7.2	331	2 O6GX85_COCCA	O6GX85 cochlilobu
36	86	7.2	893	2 O68VH4_9POTY	O68VH4 barley mild
37	86	7.2	3692	2 O5J1Q7_9NOCA	O5J1Q7 nocardia un
38	85	7.1	337	1 SYM1_STRAW	O82H01 streptomyc
39	85	7.1	671	2 O815K5_PLAUF	O815K5 plasmodium
40	85	7.1	1018	2 O6FRN6_CANGA	O6FRN6 candida gla
41	85	7.1	2194	1 GLSN_MEDSA	O03460 medicago sa
42	85	7.1	2194	1 G40360_MEDSA	O40360 medicago sa
43	84.5	7.1	361	2 O8YV34_ANASP	O8YV34 anabaena sp
44	84.5	7.1	376	2 O5YZE1_NOCFA	O5YZE1 nocardia fa
45	84	7.0	247	2 O64WS0_BACFR	O64WS0 bacteroides

## ALIGNMENTS

RESULT 1  
ID OIP5\_HUMAN STANDARD; PRT; 229 AA.  
AC O43482; O96BX7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Opa-interacting protein 5.  
GN Name=OIP5;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RP [1]  
NP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=98125741; PubMed=9466265;  
RA Williams J.M., Chen G.-C., Zhu L., Rest R.F.;  
RT "Using the yeast two-hybrid system to identify human epithelial cell  
RT proteins that bind gonococcal Opa proteins: intracellular gonococci  
RT bind pyruvate kinase via their Opa proteins and require host pyruvate  
RT for growth.";  
RT Mol. Microbiol. 27:171-186(1998).  
RN [2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RP TISSUE=uterus;  
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Scheet C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loughell N.A., Peters G.J., Abramson R.D., Millar S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L., Hulyk S.W.,  
RA Villalón D., Muly J.C., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RL -i- SUBUNIT: Binds outer membrane protein OpaP from Neisseria  
-i- gonorrhoeae.  
CC -i- INTERACTION: NBExp=3; IntAct=EBI-536879, EBI-365996;  
CC P04049:RAPI; NBExp=3; IntAct=EBI-536879, EBI-365996;  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

DR EMBL; AF025441; AAC39561.1; ALT INIT; mRNA.

DR EMBL; BC015050; AAH15050.1; -; mRNA.

DR IntAct; O43482; -.

DR Ensembl; ENSG00000104147; Homo sapiens.

DR HGNC; HGNC:20300; OLP5.

DR H-InvDB; HIX0012150; -.

DR MIM; 606020; -.

DR GO; GO:0005515; F:protein binding; TAS.

DR GO; GO:0007154; P:cell communication; TAS.

SO SEQUENCE 229 AA; 24691 MW; 0EBD406193A3106 CRC64;

Query Match 100.0%; Score 1198; DB 1; Length 229;  
Best Local Similarity 100.0%; Pred. No. 8.ee-105;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAPLRHRCATPPRPGCGTERAIDQASFTTSMNDQVVGSSPLGAGGEEP 60  
Db 1 MAAPLRHRCATPPRPGCGTERAIDQASFTTSMNDQVVGSSPLGAGGEEP 60

Qy 1 AAGPOLPSMLQPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120  
Db 61 AAGPOLPSMLQPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120

Qy 121 VGIEGSLKSTYNNLFCGSGCI PVGFHLSTHAAALRGHFCLSDDKVCYLKTKAIV 180  
Db 121 VGIEGSLKSTYNNLFCGSGCI PVGFHLSTHAAALRGHFCLSDDKVCYLKTKAIV 180

Qy 121 VGIEGSLKSTYNNLFCGSGCI PVGFHLSTHAAALRGHFCLSDDKVCYLKTKAIV 180  
Db 121 VGIEGSLKSTYNNLFCGSGCI PVGFHLSTHAAALRGHFCLSDDKVCYLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVTTHRLKSLMKIISVTPDQSKPEN 229  
Db 181 NASEMDIQNVPLSEKIAELKEKIVTTHRLKSLMKIISVTPDQSKPEN 229

Qy 181 NASEMDIQNVPLSEKIAELKEKIVTTHRLKSLMKIISVTPDQSKPEN 229  
Db 181 NASEMDIQNVPLSEKIAELKEKIVTTHRLKSLMKIISVTPDQSKPEN 229

RESULT 2

Q9CKR6\_MOUSE PRELIMINARY; PRT; 218 AA.

ID Q9CKR6\_MOUSE PRELIMINARY; PRT; 218 AA.

AC Q9CKR6; PRT; 218 AA.

DT 01-JUN-2001 (TReMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched library, clone:3110025H23 product:hypothetical protein, full insert sequence.

GN Name=2610039C10R1k;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muroidea; Muridae; Murinae; Mus.

OC NCBI\_TaxID=10090;

NC 11

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., RA Arkawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I., RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R., RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T., RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H., RA Kuhl P., Lewis S., Matsumoto Y., Nishida K., Pesole G., Quackenbush J., RA Schmitt L.M., Stebbins R., Suzuki R., Tomita M., Wagner L., Washio T., RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G., RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L., RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S., RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA The PANTOM Consortium.

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [4]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M., RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., RA Kono H., Akiyama J., Nishi K., Kikunai T., Tashiro H., Itoh M., RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M., RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Head;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., RA Arkawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C., RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M., RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T., RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshino M., RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK014084; BAB29147.1; -; mRNA.

DR Ensembl; ENSMUSG0000022978; Mus musculus.

DR MGI; MGI:1913828; 2610039C10R1k.

SO SEQUENCE 218 AA; 24447 MW; 553905BFD86C2 CRC64;

Query Match 10.2%; Score 122; DB 2; Length 218;  
Best Local Similarity 23.3%; Pred. No. 0.0049;  
Matches 48; Conservative 30; Mismatches 102; Indels 26; Gaps 6;

Qy 29 DQASFTSMNDQVVGSSPLGAGGAEPAQPOLPSMLQPERCAVFOCAQCHAVLADSV 88  
Db 15 DQASFTSMNDQVVGSSPLGAGGAEPAQPOLPSMLQPERCAVFOCAQCHAVLADSV 88

Qy 65 DSVHAMDLSR-SLGAVFSRVTNNVLEAPFLVGIEGSLKSTYNNLFCGSGCIPVGFH 147  
Db 65 DSVHAMDLSR-SLGAVFSRVTNNVLEAPFLVGIEGSLKSTYNNLFCGSGCIPVGFH 147

QY 148 LYSTHAALAAAGHFCISDDKRVCTLL-----KTKAIVNASEMDIQNVPLSEKAE 198  
DB 123 YRCFTPKNDYKRDLCFLSEVTEVSYLGSSEKQIVSKDELFNLE-----SRVEIKSIKQ 178  
QY 199 LKEKIVLTHNRKSLMKILSEVTPQ 224  
DB 179 MEEVLTALQKKLREVESKLSLAPQ 204

RESULT 3  
Q4PAR5 USTMA PRELIMINARY; PRT; 462 AA.  
ID Q4PAR5\_USTMA PRELIMINARY; PRT; 462 AA.  
AC Q4PAR5;  
DT 13-SEP-2005 (TREMBlrel. 31, Created)  
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=OM02948.1;  
OS Ustilago maydis 521.  
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;  
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.  
OX NCBI\_TaxID=237631;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=521;  
RA Birren B., Nuebaum C., Abebe A., Abouelleil A., Adekoya E.,  
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,  
RA Arachchi H., Ambusteyn J., Bachantsang P., Baldwin J., Barry A.,  
RA Bayal T., Bishereygn B., Bloom T., Bye J., Boguslavsky L.,  
RA Botowsky M., Boukhalter B., Brumache A., Butler J., Calixte N.,  
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,  
RA Calymore A., Condidne T., Cook A., Cooke P., Corum B., Cuomo C.,  
RA David R., Dawe T., Degray S., Dodge S., Dooley K., Dorje P.,  
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,  
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,  
RA Fitzgerald M., Foley K., Gage D., Galegan J., Geatin G., Gnerre S.,  
RA Guitke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hatz N.,  
RA Hagopian D., Hage B., Hall J., Hatcher B., Heller A., Higgins H.,  
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Hubby E., Iliev I.,  
RA Jaffe D., Jones C., Kanel M., Kamat A., Kamysseles M., Karlsson E.,  
RA Kells C., Kien A., Kiser P., Kodira C., Kulbokas E., Labutti K.,  
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,  
RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lueten O.,  
RA Lui A., Ma L.J., Mabit R., MacDonald J., Maclean C., Major J.,  
RA Manning J., Marbella R., Maru K., Matthews C., Maurice E.,  
RA McArthur M., McDough S., McGhee T., Meidrim J., Meneus L.,  
RA Meistrov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,  
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,  
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,  
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,  
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,  
RA Purelli S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,  
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,  
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,  
RA Shepherd B., Sherry N., Shi J., Smirnov S., Smith C., Sounez C.,  
RA Stenson K., Stone C., Stone S., Stubbs M., Talamas J., Tchinga P.,  
RA Tensing P., Tsefaye S., Theodore J., Thoulirang Y., Topham K.,  
RA Towey S., Teama T., Tsono N., Vallee D., Vassiliev H.,  
RA Vankataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
RA Vengataraman V., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,  
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,  
RA Zimmer A., Zody M., Zander E.;  
RT "The genome sequence of Ustilago maydis";  
CC Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAC0100101; EAK64120.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 462 AA; 49283 MW; EDDCFDF1757B509 CRC64;

Query Match 9.8%; Score 118; DB 2; Length 462;  
Best Local Similarity 25.9%; Pred. No. 0.03;  
Matches 64; Conservative 28; Mismatches 109; Indels 46; Gaps 10;  
QY 2 AAQPLRHSRCATPRGDFCGGTERAIDQASFTTSMEMDTQV-VKSSPLGAGIAGEBP 60  
DB 109 AAQPLSDDDMAWAPKX-----GSITTT---STRARGRGGRGRGRGRG 153  
QY 61 AA-----GROL--PSMLQPERCAVFOCAQCHAVLADSV-HLADLSRSIGAVVSRPT 110  
DB 154 ASTSTTVVHKPDLESDSEBPNNPLVFOCRCCFRLLGSLAEVARDI--DLGYVLSDVS 211  
QY 111 NNVLAEAPFVIGIE-GSLKGSYVNLFCGSCGIPVGFHLYSTHAALAAHGFCLSSDMK 169  
DB 212 EIIQDDTTESTEBGKQIGSTFPAIRACGNAVGRVYRTTRPDLDDLCDFSLVDAL 271  
QY 170 VCYLKTKRAIVNASEMDIQ-----NVPLSEKIALEKIVLTH-NRLKS 212  
DB 272 VTYGLGSNVTYRQKEDEDDQDALVDGASGKPARASRPTHTVQADPTFKVLTIKMERFPA 331  
QY 213 LMKILSE 219  
DB 332 LTIELSD 338

RESULT 4  
ID Q9CZJ6\_MOUSE PRELIMINARY; PRT; 204 AA.  
AC Q9CZJ6;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
DE Mus musculus 11 days embryo whole body cDNA, RIKEN full-length  
DE enriched library, clone=270078124 product:hypothetical protein, full  
DE insert sequence (2610039C10Rik protein).  
GN Name=2610039C10Rik;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Eucarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Whole body; DOI=10.1038/35055500;  
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1016/S0076-6879(19)03004-9;  
RX Carlini P., Hayashizaki Y.;  
RA "High-efficiency full-length cDNA cloning";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Whole body; DOI=10.1038/35055500;  
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1016/S0076-6879(19)03004-9;  
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bond H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Bataiov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuenl P., Lewis S., Mateno Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staudl P., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,  
RA Brownstein M.D., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinich S., Hill D., Hofmann M., Humé D.A., Kamuya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
RN [3]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi U., Bono H., Kondo S.,  
RN Nakairo T., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gobjohri T.,  
RN Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schirral L.M., Knapin A., Matsuda H., Batalov S., Beisel K.W.,  
RN Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.F., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frizer K.S.,  
RN Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimonod S., Guclincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RN Kanai A., Kawaji H., Kawasawa Y., Kedziarski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RN Maglocz D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numa K., Okido T., Pavan W.J., Pereira G., Pesole G.,  
RN Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RN Sadelain A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RN Velardo R., Wagner L., Wahlestedt C., Wang Y., Watnabe Y., Wells C.,  
RA Wilming L.G., Wyshnaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RN Yuan Z., Zavalon M., Zhu Y., Zimmer A., Yamanaka M., Hayatsu N.,  
RA Hirozane-Kienikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
RN Shiraki T., Waki K., Kawai J., Aizawa K., Arkhawa T., Fukuda K.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RN Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,  
RA Yaenunishi A., Yoshino M., Waterson R., Lander E.S., Rogers J.,  
RN Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
[4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/9r.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RN Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes."  
RL Genome Res. 10:1617-1630(2000).  
[5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/9r.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RN Kono H., Akiyama J., Nishi K., Kitsuai T., Taahito H., Itoh M.,  
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
RN Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kasaiwa K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RN Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
[6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX Arakawa T., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RN Arakawa T., Aono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hasegaki T., Hara A., Hayatsu N., Hiramoto K., Hiroka T., Hori F.,  
RN Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kuzihara C.,  
RN Matsumura T., Miyazaki A., Nishi K., Nomura K., Nunazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RN Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami K., Tagawa A., Takahashi F., Tanaka T.,  
RN Teijima Y., Toya T., Yamamura T., Yaenunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
[7]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Eye;

Query	Best Local Similarity	Score	DB	Length
Matches 50; Conservative 29; Mismatches 97; Indels 19; Gaps 7;	9.6%;	115.5;	2;	204;
Query Match	25.6%;	Pred. No. 0.018;		
Db	15	DSRYLRLQKWN--MSSADALGLEKERPEKKAALNP-----LVFLCAGCRPLG	64	
Qy	89	DSVLIANDLSR-SLGAVVFSRVTVNVVLEAPFLVGIEGSLGSTRYNNLFCGSCGIPVGFH	147	
Db	65	DS--LTVWAGQEDINCLILSVSCNVSDVKDEPKLSKCRDEDCGLIALYCGCSLSGYV	122	
Qy	148	LYSTHAALAIIRGHFLCSSDKMVCYLL--KTKAIVNASENDIQVNPJSEKLAELKEKIVL	205	
Db	123	YRCPTKMLDYKRDJFLCSVEAVESYTGSSSEKQIV-SEDKELFNI--ESRVEIEKSIQI	178	
Qy	206	THNRLKSLMKLISFV	220	
Db	179	MEEVLTALQKKLREV	193	
Result 5				
CU045 HUMAN				
ID CU045 HUMAN	STANDARD;	PRT;	233 AA.	
AC Q9NPY9;				
DT 16-OCT-2001	(Rel. 40, Created)			
DT 16-OCT-2001	(Rel. 40, Last sequence update)			
DT 10-MAY-2005	(Rel. 47, Last annotation update)			
DE Protein C21orf45.				
GN Name=C21orf45;				
OS Homo sapiens (human).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;				
OC Homo.				
OX NCBI_TaxID=9606;				
XP [1]				
NP NUCLEOTIDE SEQUENCE.				
RX MEDLINE=20237674; PubMed=10773462; DOI=10.1016/S0378-1119(00)00089-5;				
RA Slavov D., Hattori M., Sakaki Y., Rosenthal A., Shimizu N.,				
RA Minoshima S., Kudoh J., Yaaso M.-L., Ramser J., Reinhardt R.,				



```

RT      Refmer C.,Clancy K., Ryndtich A., Gardiner K.;
RT      "Criteria for gene identification and features of genome organization:
RT      analysis of 6.5 Mb of DNA sequence from human chromosome 21.";
RL      Gene 247:215-232(2000).
RL      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC      TISSUE=Blood;
RC      MEDLINE=22386257; PubMed=12477933; DOI=10.1073/pnas.242603699;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schult G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA      Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein W.J., Uebin T.B., Tomihataki S., Carninci P., Prange C.,
RA      Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gundaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holvik S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Pailey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA      Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallue D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
CC      -1- SIMILARITY: To S.pombe SPC970.12.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL; AF231921; AAF72945.1; -; mRNA.
DR      EMBL; BC042917; AAH42917.1; -; mRNA.
DR      Eneembl; ENSG00000159055; Homo sapiens.
DR      HGNC; HGNC:1286; C21orf45.
SQ      SEQUENCE 233 AA; 25863 MW; A07522806C4B6221 CRC64;

Query Match      9.4%; Score 113; DB 1; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.036;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY      31 ASFTTSMMDNQVYVSGSPPLGPAGAGAEPPAAGPOLPMLQPERCAVPOCAQCHAVLADS 90
DB      49 ASMSMSMEDSDSV---ADMERAQL--EEEAAMAE-----ERLVTLSCGCRRLPDS 95
QY      91 VHLAWDLR-SIGAVVFRVYTNVNVLLAPFLVIGSLGSLKSTYNNLLFCGSGCIPVGFHY 149
DB      96 --LSVVAQEDPTNCLLRCSVSNVSVDRKQKLSKREKENGCVLRTLCAGCSLNLGYVR 153
QY      150 STHAALLALRGHFCILSSDRMCYLL--KTKAIVNASENDIQVPLSEKIAELKEIVYLTH 207
DB      154 CTPKMLDYKRDIFCLCSVEAIESYVLSGSEKQIV--SEDKELFVL--ESRVIETKSLTOME 209
QY      208 NRTKSLMKTLSE 219
DB      210 DVLKALQMKLWE 221

RESULT 6
OS4220_HUMAN
ID OS4220_HUMAN PRELIMINARY; PRT; 233 AA.
AC OS4220;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Iast sequence update)
DE 13-SEP-2005 (TRENBLREL. 31, Iast annotation update)
DE Hypothetical protein FLJ90800 (FAPPI-associated protein 1).
GN Name=FAPPI.
OS Homo sapiens (Human).

```

[illegible]

RA Taenzer S., Galgoczy P., Platzner M., Scharfe M., Nordstiek G.,  
 RA Blocker H., Hellmann I., Khatovich P., Prabo S., Reinhardt R.,  
 RA Zhang H.-J., Zhang X.-L., Zhu G.-F., Wang B.-F., Fu G., Ren S.-X.,  
 RA Zhang G.-P., Chen Z., Lee Y.-S., Cheong J.-E., Choi S.-H., Wu K.-M.,  
 RA Liu T.-T., Hsieh K.-J., Tsai S.-F., Kim C.-G., Oota S., Kitano T.,  
 RA Kohara Y., Saitou N., Park H.-S., Wang S.-Y., Yasuo M.-L., Sakaki Y.,  
 RT "DNA sequence and comparative analysis of chimpanzee chromosome 22.";  
 RL Nature 429:382-388(2004).  
 CC -1- SIMILARITY: To S.pombe SPC970.12.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 DR EMBL, AL954205; CAH18576.1; -; Genomic DNA.  
 SQ SEQUENCE 232 AA; 25832 MW; 654988BE369245E7 CRC64;  
 Query Match 9.2%; Score 110.5; DB 1; Length 232;  
 Best Local Similarity 27.1%; Pred. No. 0.065;  
 Matches 52; Conservative 29; Mismatches 88; Indels 23; Gaps 8;  
 Oy 31 ASFTSMWDPTOVVYSSPLGPAAGAEPPAAGPOLPSMLQPERCAVFOCAQCHAVLADS 90  
 Db 49 ASMSMSMEDASV-----ADMERARL--EEAAA-----EERPVFLGCGCRRPLGDS 94  
 Oy 91 VHLAMDLSR-SLGAVFPSRVNNVYL-LEAPFLVIGSLKSTYNLLFCGSGCIPVGHLY 149  
 Db 95 --LSWVAGSDPTNCTILRCVSCNVSDKEQKLSKREKENGCVLETLCCAGSLNMGVYR 152  
 Oy 150 STHAALALRGHFLCSSDKMVCYL--KTKAIVNASEMDIONVPLSEKIALKKEIVLTH 207  
 Db 153 CTPKRLDYKRLFCILSVLAISVYLSGSEKQIV-SEDKELFNL---ESRVEIKESLTQME 208  
 Oy 208 NRLKSLMKLISE 219  
 Db 209 DVLKRLQMKLME 220  
 RESULT 8  
 OSBLBI BRARE  
 ID OSBLBI BRARE PRELIMINARY; PRT; 207 AA.  
 AC 10-MAY-2005 (Tremblrel. 30, Created)  
 DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)  
 DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)  
 GN LOC553502 protein (Fragment).  
 GN Name:LOC553502;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Singapore local strain; TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Dichtchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stopleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Wuzny D.M., Sodergren B.U., Lu X., Gibbs R.A.,  
 RA Paley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skálaka U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Singapore local strain; TISSUE=Embryo;  
 RG NIH MGC Project;  
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC090521; AA090521.1; -; mRNA.  
 DR InterPro; IPR000074; ApoA1\_A4\_E.  
 DR InterPro; IPR009090; D\_amino\_dept\_C.  
 FT NON TER 1  
 SQ SEQUENCE 207 AA; 22941 MW; AFIADACC375A28C CRC64;  
 Query Match 9.1%; Score 109; DB 2; Length 207;  
 Best Local Similarity 25.5%; Pred. No. 0.077;  
 Matches 37; Conservative 28; Mismatches 66; Indels 14; Gaps 6;  
 Oy 76 AVFOCAQCHAVLADSVHLAMDLS-RSLGAVFPSRVNNVYL-EAPFLVIGSLKSTYN 133  
 Db 54 AVFWCGCKPLIGDS--LSWAGSDDENNQIMLRITDNIYVKGKPFVGTREKLGCLVNV 111  
 Oy 134 LIFCGSGCIPVGHLYSTHAALALRGHFLCSSDKMVCYLKTKAIVNASEMDIONVPLS 193  
 Db 112 LT-CRCCGSELGVMYISTPKLDOCKSLFCFNNENIESYVGSBG-QQMELPREDKPV 169  
 Oy 194 -----EKIABLKKEIVLTHNL 210  
 Db 170 LEYQDVHQQMETITSLAVIIGRL 194  
 RESULT 9  
 OS4H16 DICDI  
 ID OS4H16 DICDI PRELIMINARY; PRT; 532 AA.  
 AC OS4H16;  
 DT 13-SEP-2005 (Tremblrel. 31, Created)  
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)  
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=DD80189417;  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.  
 OC NCBI\_TaxID=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sungang B., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Kontorov B.A., Rivero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill B., Bason N.,  
 RA Farbrother P., Desany B., Just E., Morit T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wadrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Loulsgaard H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Bartell B., Kuspa A.,  
 RT "The genome of the social amoeba Dictyostelium discoideum.";  
 RL Nature 0-0-0(2005)  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL, AAF01000207; EAL62738.1; -; Genomic DNA.  
 KW Hypothetical protein.

SEQUENCE 532 AA; 60715 MW; BA75AB8CD7C1D5C CRC64;

Query Match 8.9%; Score 107; DB 2; Length 532;  
Best Local Similarity 22.9%; Pred. No. 0.4;  
Matches 38; Conservative 29; Mismatches 75; Indels 24; Gaps 5;

QY 77 VFQCAQCHAVLADSVHLAMDLSRSLGAVVFSRVTVNNVLAEPFLVIGISLKGST----- 131  
DB 205 IFSCLSCKPTIDSTLI---VNSKSHCDVMTLTYTPFKKLGSGSVILDDQKSTRGRT 261  
QY 132 ---VVLTFPGSGCGIPVGFHLVSTHAALALRGHFLCSDRWKVCYLL-KTKAIVNASEMDI 187  
DB 262 DCTVSLFYCYCNSPLGKRTKNTKSQFKILNNHFPVVDISITTYISVGGNMNIIVEDKRI 321  
QY 188 QNVPLS-----EKIAELKEKIVLTNRLKSLMKLISEVTPD 223  
DB 322 ---VPLSFGNDKLTQKHETMKNNKNEKINFLDDQLLSFVNSIKKALPE 365

RESULT 10

YCZC SCHPO STANDARD; PRT; 155 AA.  
AC 09P802;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-MAY-2005 (Rel. 47, Last annotation update)  
DE Hypothetical protein C970.12 in chromosome III.  
GN ORFNames=SPCC970.12;  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
NCBI\_Taxid=4896;

NCBI\_Taxid=4896;

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=972;  
RA MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,  
RA Sgouras J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Frazer A.,  
RA Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K.L., Murphy L.D., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowles E.,  
RA Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M.N., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J.R., Voicikert G., Aert R., Robben J., Grymoprez B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Duesterhoeft A., Fritz C., Holzer E., Moestl D.,  
RA Hilbert H., Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R.,  
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambut R., Purrelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forzberg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shparovskii G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).

-1- SIMILARITY: To human C21orf45.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC EMBL: AL031530; CAB72327.1; -; Genomic\_DNA.

DR GeneDB\_Spombe; SPC970.12; -.

DR GO: GO:0005515; F:protein binding; IPI.

DR GO: GO:0031055; P:chromatin remodeling; at centromere; IMP.

DR GO: GO:0007059; P:chromosome segregation; IMP.

DR GO: GO:0016575; P:histone deacetylation; IMP.

DR GO: GO:0031066; P:regulation of histone deacetylation at cent. .; IMP.

KW Complete proteome; Hypothetical protein.

SEQUENCE 155 AA; 17874 MW; A7ABD0F93760B08 CRC64;

Query Match 8.8%; Score 105; DB 1; Length 155;  
Best Local Similarity 30.1%; Pred. No. 0.13;  
Matches 41; Conservative 18; Mismatches 61; Indels 16; Gaps 6;  
QY 71 OPERAVFOCAQCHAVLADSVHLAMDLSRSLGAVVFSRVTVNNVLAEPFLVIGISLKGST 128  
DB 19 QP---SVFOCKCKCFQIVGDS--NAVISHRETLSTLSDAENSRAVEDTFKRSDDGL-- 71  
QY 129 GSTVNLTFPGSGCGIPVGFHLVSTHAALALRGHFLCSDRWKVCYLLKTKAIVNASEMDIO 188  
DB 72 -CVSELSCTRCNEBVGKRVNSTPIYLDIRDMYTFMSMDKLAQAYOLGNKT-VNPEGLTRY 129  
QY 189 NVPLSEKIAELKEKIV 204  
DB 130 QVDL-----EMREDII 140

RESULT 11

Q40626 LEIMA PRELIMINARY; PRT; 1556 AA.  
AC Q40626;  
DT 13-SEP-2005 (TRENBLrel. 31, Created)  
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=LmjF31.2350;  
OS Leishmania major.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
NCBI\_Taxid=5664;

NCBI\_Taxid=5664;

NUCLEOTIDE SEQUENCE.

RC STRAIN=Friedlin;  
RA Peacock C.S., Murphy L., Ivens A.C., Berriam M., Blackwell J.,  
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,  
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;  
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
RL EMBL: C7005268; CAJ08424.1; -; Genomic\_DNA.  
KW Hypothetical protein.

SEQUENCE 1556 AA; 170195 MW; F6F7CDDEB8DCA59 CRC64;

Query Match 7.8%; Score 94; DB 2; Length 1556;  
Best Local Similarity 24.3%; Pred. No. 27;

Matches 45; Conservative 22; Mismatches 50; Indels 68; Gaps 9;

QY 22 GGERALIDQASFTTSM-----WPTQVVKGS-----SPLQAGIAGAEPPAAGPO 65  
DB 833 GGCELMQTDAGSIEVESPILAFYDDVDHRAQDVQKARREPPDVGAGGEDAAG-- 890  
QY 66 LPSTWLOPERCAVFOCAQCHAVLADSV-----HLAMDLSRLGAVVFSRVTVNNVLEAP 118  
DB 891 -----ARAAIGE-QDFGFLNDVVAAPSTYSHMLCRVLEIRFARLMLILDLVLLQAP 941  
QY 119 FLVVGISLKGSTVNLTFPGSGCGIPVGFHLVSTHAALALRGHFLCSDRWKVC---YLL 174  
DB 942 VL-----TRY-----VWYHLYK-----HFCLSADKASCVIQGMILL 971

QY 175 KTKAI 179

DB 972 KMLAV 976

RESULT 12

Q8S611 ORYSA

ID Q8S611\_ORYSA PRELIMINARY; PRT; 361 AA.

AC Q8S611;

DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DE 01-FEB-2005 (Tremblrel. 29, Last annotation update)  
DE Putative gypay-type retrotransposon protein.  
GN ORFNames=OSJNB0096822.5;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriocarpaceae; Oryzae; Oryza.  
NCBI\_TaxID=39947;  
RN Nucleotide Sequence.  
RP Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Sasaki C., Henry D., Oates R., Simmons U.,  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN Nucleotide Sequence.  
RP The Rice Chromosome 10 Sequencing Consortium;  
RT "In-depth view of structure, activity, and evolution of rice  
RT chromosome 10."  
RT Science 300:1566-1569(2003).  
RN Nucleotide Sequence.  
RP Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.,  
RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AC099400; AAL91599.1; -; Genomic DNA.  
DR EMBL; AE017047; AAP51763.1; -; Genomic DNA.  
DR Gramine; Q85611; -.  
SQ SEQUENCE 361 AA; 38804 MW; F883BB3EB8B8FF45 CRC64;

Query Match 7.7%; Score 92; DB 2; Length 361;

Best Local Similarity 25.3%; Pred. No. 6.4;  
Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;

QY 2 AAOPLRHSRCATPRGDFCGTERAIDQASFTTSMENDTVVKKSSPLGAGIAGEEPA 61  
DB 49 APTPLPFRRAVRARAKDGGSGTSSAPVAST-----DVVVVPSRRATSG-PASDPV 102  
QY 62 AGPOLP-----SW-LQPERCAVFOCAQCHAVLADSVHLAMDLSLSLGAVFESRVNTNVL 115  
DB 103 AGRGSPAVLVSWELQVEMGRLLBGA--RVIGREIARGLGHEHM-----SELGNN--- 152  
QY 116 EAPPLVLEGSLKSTYML--FCGSGCI-----PVGFILYSTHALLAL-----RGH 161  
DB 153 -----LSEIRGSLR-VTYTGLHQLACKGCKSTIPANPEFSLTSLAELAMEBIPSKH 207  
QY 162 FCLSDSKM-----VCYLTKAIVNASEMDIQNV-----PLSEKIAELKEKI 203  
DB 208 AARIGESNRIYIGACHILACVRLAH-PELDRLRIIDOGKASDPARKDVMEEVDLGKSV 266  
QY 204 V 204  
DB 267 L 267  
RESULT 13  
Q9CH87 LACIA PRELIMINARY; PRT; 695 AA.  
ID Q9CH87 LACIA PRELIMINARY; PRT; 695 AA.  
AC Q9CH87;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Copper-potassium transporting ATPase B.  
GN Name=COB; OrderedLocNames=LI0851;  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
NCBI\_TaxID=1360;  
RN Nucleotide Sequence.  
RP Nucleotide Sequence.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr-1697R;  
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,  
RA Weissenbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis IL1403."  
RL Genome Res. 11:731-753(2001).  
DR EMBL; AE006319; AAK04949.1; -; Genomic DNA.  
DR PIR; C86731; C86731.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0004008; F:copper-exporting ATPase activity; IEA.  
DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid. . .; IEA.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0046873; F:metal ion transporter activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0030001; P:metal ion transport; IEA.  
DR GO; GO:0015992; P:proton transport; IEA.  
DR InterPro; IPR006403; ATPase-IB1\_Cu.  
DR InterPro; IPR006416; ATPase-IB\_hvy.  
DR InterPro; IPR001757; ATPase E1-E2.  
DR InterPro; IPR005834; Dehal like hydro.  
DR InterPro; IPR008250; E1-E2\_ATPase\_reg.  
DR Pfam; PF00122; E1-E2\_ATPase; 1.  
DR Pfam; PF00702; Hydrolyase; 1.  
DR PRINTS; PR00119; CATATPASE.  
DR TIGRFAMs; TIGR01511; ATPase-IB1\_Cu; 1.  
DR TIGRFAMs; TIGR01525; ATPase-IB\_hvy; 1.  
DR TIGRFAMs; TIGR01494; ATPase-P-type; 2.  
DR PROSITE; PS00154; ATPASE\_E1\_E2; UNKNOWN\_1.  
KM Complete proteome.  
SQ SEQUENCE 695 AA; 75981 MW; 1B4947C3C2A0FA0F CRC64;

Query Match 7.7%; Score 92; DB 2; Length 695;

Best Local Similarity 25.3%; Pred. No. 15;  
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;

QY 101 LGAVFSRVNNTNVLAEPLVLEGSLKSTYMLFCGSGCIPVGFILYSTHALLALRG 160  
DB 101 LGTIIF-----FYSGRPFSGAKGELSKRPAMMMLITMGITVAV-AVSYATIMSLNG 153  
QY 161 HF-----CLSDKMCVYLTKAIVNASE-MDI-----QNVPLSE-K 195  
DB 154 HMGNNFWFELATLVIMLIGLIEKMAKNGKDALKDLASLVPRKAILKSGKDELSELK 213  
QY 196 IAELEKIVLTHNRKSLMKILSEVTPDOS 225  
DB 214 VGDL-LLVKENEKIPADGILSLVALVDES 241  
RESULT 14  
Q69KE0 ORYZA PRELIMINARY; PRT; 628 AA.  
ID Q69KE0 ORYZA PRELIMINARY; PRT; 628 AA.  
AC Q69KE0;  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DE 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Myosin heavy chain-like.  
GN Name=P0023E10.17;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriocarpaceae; Oryzae; Oryza.  
NCBI\_TaxID=39947;  
RN Nucleotide Sequence.  
RP Nucleotide Sequence.  
RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.,  
RT "Oryza sativa nupprobare (GA3) genomic DNA, chromosome 9, PAC  
RT clone: P0023E10.17".  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP005934; BAD36600.1; -; Genomic DNA.  
DR Gramine; Q69KE0; -.  
DR InterPro; IPR005829; Sug transporter.  
DR PROSITE; PS00216; SUGAR TRANSPORT 1; UNKNOWN\_1.  
SQ SEQUENCE 628 AA; 68281 MW; 3330B108C6061027 CRC64;



**This Page Blank (uspto)**